

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE GEOGRAFÍA E HISTORIA



**MÁSTER UNIVERSITARIO EN
TECNOLOGÍAS DE LA INFORMACIÓN GEOGRÁFICA**

TRABAJO FIN DE MÁSTER

CURSO 2019-2020

**MODELOS DE ONDULACIÓN GEOIDAL PARA LA RED TOPOGRÁFICA DE
MADRID A TRAVÉS DE TÉCNICAS GEOESTADÍSTICAS DISEÑADAS A
PARTIR DE SOFTWARE LIBRE Y COMERCIAL.**

**GEOID UNDULATION MODELS FOR MADRID TOPOGRAPHIC NETWORK
THROUGH GEOSTATISTICAL TECHNIQUES DESIGNED FROM FREE AND
COMMERCIAL SOFTWARE.**

JORGE RAINIERO ESTRELLA SALVADOR

CONVOCATORIA: OCTUBRE 2020

TUTORES:

Luis Miguel Tanarro García
Departamento de Geografía (Fac. Geografía e Historia)

María Teresa Benavent Merchán
Unidad Departamental de Astronomía y Geodesia (Fac. de Matemáticas)

RESUMEN

Para la Geodesia, la ondulación geoidal se define como la diferencia entre el Geoide y un elipsoide de referencia. Esta variable permite corregir una altura geometría a una altura física o real. Dicha corrección es de importancia para el diseño y ejecución de diversos proyectos de ingeniería a nivel mundial, los cuales necesitan puntos georreferenciados que consideren la forma real de la tierra (referidos al geoide) y la influencia de sus campos gravitatorios. Bajo este criterio, el municipio de Madrid es una comunidad que constantemente está creciendo y cambiando por su gran dinámica económica, social, turística, entre otros aspectos. Por lo que es una ciudad que requiere de infraestructura que le permita satisfacer dichas necesidades.

De modo que, el presente trabajo de fin de Master, tuvo como objetivo generar a través de técnicas geoestadísticas dos modelos de la variable ondulación geoidal -en software libre y comercial- para el municipio de Madrid, utilizando como base su red topográfica. Estos modelos fueron diseñados a partir de un análisis exploratorio de datos, análisis estructural de un variograma experimental y el uso de kriging como método de interpolación. Posteriormente los modelos resultantes fueron cotejados entre sí y con el modelo español EGM08-REDNAP, con el fin de determinar la calidad de los mismos.

Palabras Claves: Ondulación Geoidal, SIG, Kriging, Geoestadística, software libre.

ABSTRACT

For Geodesy, the geoid undulation is defined as the difference between the Geoid and a reference ellipsoid. This variable allows correcting a geometry height to a physical or real height. This correction is important for the design and execution of different engineering projects worldwide, which needs georeferenced points that considers the real shape of the earth (referred to the geoid) and the influence of its gravitational fields. Under this criterion, the municipality of Madrid is a community that is constantly growing and changing due to its great economic, social and tourist dynamics, among other aspects. So, it is a city that requires infrastructure that allows it to satisfy these needs.

Therefore, the objective of this final Master's work was to generate, through geostatistical techniques, two models of the geoidal undulation variable -in free and commercial software- for the municipality of Madrid, using its topographic network as base. These models were designed from an exploratory data analysis, structural analysis of an experimental variogram, and the use of kriging as an interpolation method. Subsequently, the resulting models were compared with each other and with the Spanish EGM08-REDNAP model, in order to determine their quality.

Keywords: Geoid undulation, GIS, Kriging, Geostatistics, free software.

DEDICATORIA

“Tu enfoque determina tú realidad”

-Qui Gon Jinn

Este trabajo se lo dedico a mis padres y a mis abuelitos, ya que sin ellos nada de esto hubiera sido posible. Este último año ha sido duro, lleno de retos y sobre todo muchos cambios, los cuales me han permitido valorar más a mi familia (a pesar que todavía no les extraño) y aquellas pequeñas cosas que nos provocan felicidad sin siquiera darnos cuenta.

Como lo mencione en mi TFG, ser paciente es una virtud que todavía no adquiero y me cuesta tanto. Sin embargo, hay un camino muy largo que todavía debo recorrer y por ello debo aprender a serlo, ya que terminar mis estudios en la UCM, solo ha sido el primer paso.

Jorge Rainiero Estrella Salvador

AGRADECIMIENTOS

Primero quiero agradecer a Dios y a la Fuerza por soportarme, sé que no soy su mejor creación, pero ahí seguimos (para aquellos lectores que no me conocen, tengo un carácter muy irónico y sarcástico, así que pido disculpas de antemano). También agradezco de sobre manera a mis padres y abuelitos, ya que literalmente sin ellos no hubiera podido estar en Madrid para terminar mi Master y seguir con mis planes que espero pronto se materialicen.

De igual manera deseo agradecer a mis tutores Luis Miguel Tanarro y Maite Benavent, quienes han sido una guía y ayuda durante todo este trabajo. No obstante, quiero extender un doble agradecimiento a Luis Miguel Tanarro, ya que no solo es mi tutor en este TFM, sino también ha sido un maestro y amigo que, desde mi llegada, me brindo su ayuda desinteresadamente (Ps: Me disculpo por todos los correos enviados).

Finalmente, para terminar con estos agradecimientos (no quiero que sea tan largo como los del TFG) los dividiré en dos grupos. El primero, será aquellos amigos que me dieron su apoyo “emocional” y/o me han soportado desde siempre, no solo este último año. Y el segundo grupo, que será aquellos amigos que ayudaron de una u otra forma con este trabajo de fin de Master, es decir los que de verdad aportaron con algo tangible. Una vez especificado esto, procederé a enlistar:

Primer grupo: Mis mejores amigos Miky, Negro y Andrés (la caterva de siempre), Lalo, Vane y Pame, quizá hay más, pero no recuerdo así que quedarán como “otros” o “etc”.

Segundo grupo: A mis amigos y compañeros del IGM (Instituto Geográfico Militar ecuatoriano) Alfonso Morillo, Christian Barahona, Alejandro Martines (tu no aportaste en nada, pero eres pana) y José Luis Carrión. Ah es verdad casi me olvido, Miky y mi hermana Samantha ustedes también están en este grupo.

Eso sería todo, así que gracias totales.

Que la fuerza los acompañe.

Jorge Rainiero Estrella Salvador

ÍNDICE DE CONTENIDO

| | |
|---|-----------|
| CAPÍTULO I. INTRODUCCIÓN | 10 |
| 1.1 OBJETIVOS..... | 11 |
| 1.1.1 OBJETIVOS GENERALES | 11 |
| 1.1.2 OBJETIVOS ESPECÍFICOS | 11 |
| 1.2 JUSTIFICACIÓN | 11 |
| CAPÍTULO II. FUNDAMENTOS TEÓRICOS | 13 |
| 2.1 SISTEMAS GEODÉSICOS DE REFERENCIA | 13 |
| 2.1.1 SISTEMA DE REFERENCIA | 13 |
| 2.1.2 MARCO DE REFERENCIA | 13 |
| 2.2 SUPERFICIES DE REFERENCIA | 13 |
| 2.2.1 GEOIDE | 13 |
| 2.2.2 ELIPSOIDE..... | 14 |
| 2.3 SISTEMA DE COORDENADAS | 14 |
| 2.3.1 COORDENADAS CARTESIANAS | 14 |
| 2.3.2 COORDENADAS GEODÉSICAS | 15 |
| 2.3.3 COORDENADAS PLANAS (UTM)..... | 16 |
| 2.4 ALTURAS DE REFERENCIA..... | 17 |
| 2.4.1 ALTURAS ELIPSOIDALES | 17 |
| 2.4.2 ALTURAS ORTOMÉTRICAS | 18 |
| 2.5 ONDULACIÓN GEOIDAL | 20 |
| 2.5.1 MODELO DE ONDULACIÓN GEOIDAL DE ESPAÑA (EGM08-REDNAP)..... | 21 |
| CAPITULO III. METODOLOGÍA..... | 23 |
| 3.1 INFORMACIÓN Y DATOS DE PARTIDA..... | 23 |
| 3.2 OBTENCIÓN DE LA ONDULACIÓN GEOIDAL DE LA RED TOPOGRÁFICA MUNICIPAL DE MADRID (RTM)..... | 24 |
| 3.3 ANÁLISIS Y DEPURACIÓN DE LOS VALORES DE ONDULACIÓN GEOIDAL DE LA RTM..... | 25 |
| 3.4 MÉTODOS DE INTERPOLACIÓN (KRIGING) Y ESTADÍSTICA ESPACIAL PARA ESTIMAR LA ONDULACION GEOIDAL EN LA SUPERFICIE DEL MUNICIPIO DE MADRID..... | 27 |
| 3.4.1. MÉTODOS DE INTERPOLACIÓN | 27 |
| 3.4.1.1. Métodos determinísticos | 27 |
| 3.4.1.2 métodos estocásticos | 28 |
| 3.4.2 ESTADÍSTICA ESPACIAL | 28 |
| 3.4.2.1 Geoestadística..... | 28 |
| 3.4.2.2 Geoestadística y la teoría de las variables regionalizadas | 28 |
| 3.4.2.3 Estacionaridad | 29 |

| | |
|---|-----------|
| 3.4.2.3.1 Estacionariedad Intrínseca o Débil | 30 |
| 3.4.3 ANÁLISIS EXPLORATORIO DE DATOS | 30 |
| 3.4.4 ANÁLISIS DE ESTRUCTURA | 31 |
| 3.4.4.1 Variograma experimental..... | 31 |
| 3.4.4.2 Covariograma | 31 |
| 3.4.4.3 Modelos teóricos de la semivarianza..... | 32 |
| 3.4.4.3.1 Modelo Exponencial | 33 |
| 3.4.4.3.2 Modelo Gaussiano..... | 33 |
| 3.4.4.3.3. Modelo Esférico | 33 |
| 3.4.4.3.4 Modelo efecto pepita puro..... | 34 |
| 3.4.5 PREDICCIÓN ESPACIAL KRIGING | 34 |
| 3.4.5.1 Kriging Ordinario..... | 34 |
| 3.4.5.2 Kriging Simple | 35 |
| 3.4.5.3 Kriging Universal..... | 35 |
| 3.4.6 VALIDACIÓN KRIGING | 36 |
| CAPITULO IV. RESULTADOS | 37 |
| 4.1 MODELO KRIGING USANDO EL SOFTWARE ARCGIS..... | 37 |
| 4.1.1 ANÁLISIS EXPLORATORIO DE DATOS | 37 |
| 4.1.2 VARIOGRAMA EXPERIMENTAL..... | 39 |
| 4.1.3 VARIOGRAMA TEÓRICO | 41 |
| 4.1.4 VALIDACIÓN CRUZADA INTERNA..... | 43 |
| 4.1.5 SUPERFICIE DE PREDICCIÓN | 44 |
| 4.2 MODELO KRIGING USANDO EL SOFTWARE R-STUDIO | 46 |
| 4.2.1 ANÁLISIS EXPLORATORIO DE DATOS | 46 |
| 4.2.2 VARIOGRAMA EXPERIMENTAL..... | 49 |
| 4.2.3 NUBE VARIOGRÁFICA..... | 50 |
| 4.2.4 VARIOGRAMA TEÓRICO | 51 |
| 4.2.5 VALIDACIÓN CRUZADA INTERNA..... | 53 |
| 4.2.6 SUPERFICIE DE PREDICCIÓN | 54 |
| 4.3 VALIDACIÓN CRUZADA EXTERNA: COMPARACIÓN DE LOS MODELOS KRIGING ARCGIS Y R-STUDIO CON EL MODELO EGM08 – REDNAP (IGN)..... | 55 |
| 4.4 PRECISIÓN DE MODELOS..... | 56 |
| 4.5 DISTRIBUCIÓN DE ERRORES | 57 |
| 4.5.1 TENDENCIA DE RESIDUALES DE LA SUPERFICIE KRIGING MODELADA EN ARCGIS | 57 |
| 4.5.2 TENDENCIA DE RESIDUALES DE LA SUPERFICIE KRIGING MODELADA EN R- STUDIO | 59 |
| 4.6 ESTABILIDAD DE LOS MODELOS | 61 |

| | |
|---|-----------|
| CAPITULO V. CONCLUSIONES Y RECOMENDACIONES | 65 |
| 5.1 CONCLUSIONES..... | 65 |
| 5.2 RECOMENDACIONES..... | 66 |
| BIBLIOGRAFÍA..... | 67 |
| ANEXOS | 70 |

ÍNDICE DE FIGURAS

| | |
|---|----|
| Figura 1: Sistema de coordenadas cartesianas | 15 |
| Figura 2: Sistema de coordenadas geodésicas y cartesianas | 16 |
| Figura 3: Distorsiones de proyección cilíndrica transversa | 17 |
| Figura 4: Nivelación geométrica..... | 18 |
| Figura 5: Diferencias entre superficies de nivel..... | 19 |
| Figura 6: Alturas de referencia..... | 21 |
| Figura 7: Ondulación geoidal RTM por cuartiles, sin depurar. | 25 |
| Figura 8: Histograma y descripción de N con datos crudos, en ArcGIS. | 26 |
| Figura 9: Diagrama de caja y bigote datos crudos. | 26 |
| Figura 10: Selección de datos atípicos de una desviación estándar positiva. | 26 |
| Figura 11: Diagrama de caja y bigote de datos depurados..... | 27 |
| Figura 12: Dos puntos a una distancia vectorial h | 29 |
| Figura 13: Disposición de muestras separadas por h | 31 |
| Figura 14: Estructura general de un semivariograma. SEMEXP corresponde al semivariograma experimental y MODELO al ajuste de un modelo teórico. | 32 |
| Figura 15: Comparación de los modelos teóricos exponencial, esférico y gaussiano. | 33 |
| Figura 16: Modelo pepita puro..... | 34 |
| Figura 17: Histograma de ondulación geoidal con datos depurados, en ArcGIS. | 37 |
| Figura 18: Gráfico de normalidad de N. | 38 |
| Figura 19: Mapa de Voronoi de la media de N. | 38 |
| Figura 20: Estacionaridad de N en proyección 3D. Verde: Proyección XZ, Azul: Proyección YZ, Rojo: Proyección XY, Amarillo: Proyección XYZ. | 39 |
| Figura 21: Componentes del vector h. θ es el ángulo de coordenadas polares | 40 |
| Figura 22: Variograma con Lag de 200 a 250 m con ángulo de barrido de 45° y de direcciones 0°, 90°, 45° y 135° respectivamente..... | 41 |
| Figura 23: Modelado de variograma teórico de N en ArcGIS | 42 |
| Figura 24: Matriz de búsqueda definida por ArcGIS | 43 |
| Figura 25: Validación cruzada interna por ArcGIS | 44 |
| Figura 26: Superficie de estimación de N para el municipio de Madrid, usando ArcGIS (ver el mapa en detalle en el Anexo 4) | 45 |
| Figura 27: Superficie de estimación de la desviación estándar de N para el municipio de Madrid, usando ArcGIS (ver el mapa en detalle en el Anexo 5)..... | 45 |
| Figura 28: Histograma Ondulación Geoidal – Madrid, en R..... | 47 |
| Figura 29: Diagrama de caja de N, en R | 47 |
| Figura 30: Estacionariedad de la variable en X..... | 48 |
| Figura 31: Estacionariedad de la variable en Y..... | 48 |
| Figura 32: Distribución espacial de datos | 49 |

| | |
|--|----|
| Figura 33: Variogramas direccionales generado en R | 49 |
| Figura 34: Variograma experimental omnidireccional generado en R | 50 |
| Figura 35: Nube variográfica generada en R | 51 |
| Figura 36: Ajuste de variograma teórico de N en R..... | 51 |
| Figura 37: Ajuste de variograma teórico con el mismo modelo anidado de ArcGIS | 52 |
| Figura 38: Ajuste de variograma teórico solo con modelo esférico..... | 52 |
| Figura 39: Error de validación cruzada interna por R..... | 53 |
| Figura 40: Superficie de estimación de N para la ciudad de Madrid, usando R (ver mapa en el Anexo 6) | 54 |
| Figura 41: Índice de Moran para validación cruzada interna (modelo ArcGIS)..... | 58 |
| Figura 42: Índice de Moran para validación cruzada externa (modelo ArcGIS) | 59 |
| Figura 43: Índice de Moran para validación cruzada interna (modelo R) | 60 |
| Figura 44: Índice de Moran para validación cruzada externa (modelo R)..... | 60 |
| Figura 45: Zonas con heteroscedasticidad sobre superficie de estimación kriging generada con ArcGIS, para validación cruzada interna (ver mapa detallado en el Anexo 8) | 61 |
| Figura 46: Zonas con heteroscedasticidad sobre superficie de estimación kriging generada con R, para validación cruzada interna (ver mapa detallado en el Anexo 9) | 62 |
| Figura 47: Zonas con heteroscedasticidad sobre superficie de estimación kriging generada con ArcGIS, para validación cruzada externa (ver mapa detallado en el Anexo 10)..... | 62 |
| Figura 48: Zonas con heteroscedasticidad sobre superficie de estimación kriging generada con R, para validación cruzada externa (ver mapa detallado en el Anexo 11)..... | 63 |

ÍNDICE DE TABLAS

| | |
|--|----|
| Tabla 1: Estructura de la base de datos de la RTM. | 24 |
| Tabla 2: Parámetros del ajuste del variograma teórico al experimental..... | 42 |
| Tabla 3: Parámetros de Estructuras Básicas - ArcGIS | 43 |
| Tabla 4: Resumen estadístico de la variable N, generado por R | 47 |
| Tabla 5: Parámetros de Estructuras Básicas - R..... | 52 |
| Tabla 6: Validación cruzada externa de modelos Kriging VS modelo IGN | 55 |
| Tabla 7: Resumen de precisiones alcanzadas por cada modelo analizado | 56 |
| Tabla 8: Resumen de estabilidad de residuales en modelos analizados | 61 |
| Tabla 9: Vértices con residuales altos en función a su entorno, localizados por Moran.... | 64 |

CAPÍTULO I. INTRODUCCIÓN

La Geodesia es la ciencia que estudia la forma de la tierra y la influencia de sus campos gravitatorios (Drewes, 2014). Para esta ciencia, dicha forma es representada por el geoide, superficie que es imposible de representar a través de modelos matemáticos exactos por su irregularidad geográfica, influencia de los campos gravitatorios, densidad terrestre y otras variables. Para su estudio, sin embargo, es esencial definir una superficie regular, representada por figuras geométricas y perfectamente definibles por la matemática (Seeber, 2003). Esta superficie regular, referida al geoide, se denomina elipsoide.

En la actualidad los Sistemas de Posicionamiento Global (acrónimo en inglés GNSS, Global Navigation Satellite System) pueden obtener la posición de un punto sobre la superficie terrestre referida a un sistema de referencia tridimensional geocéntrico (coordenadas cartesianas X, Y, Z). No obstante, dicho punto antes de ser usado para cualquier aplicación, debe ser transformado de coordenadas cartesianas a geodésicas: es decir cambiarán de $[X, Y, Z]$ a $[\lambda, \phi, he]$ (Hernández, S/F) (se definirán detalladamente en apartados posteriores). Sin embargo, la componente he necesita de otra transformación: reducir la altura elipsoidal a una ortométrica (H), cada una referida al elipsoide y geoide respectivamente. Esta primera, altura elipsoidal, no es empleada para ningún tipo de aplicación de ingeniería, navegación, geodinámica, etc. debido a que es una magnitud que tiene sentido geométrico, pero no físico, al no tener en cuenta la fuerza de la gravedad terrestre, mientras que la altura ortométrica lo hace (Buill Pozuelo y Núñez, 2005).

La relación existente entre el geoide y elipsoide está definida por medio de una variable denominada "ondulación del geoide" (Buill Pozuelo y Núñez, 2005). Gracias a esta variable es posible reducir una altura con un significado geométrico (medida sobre un modelo matemático terrestre, el elipsoide) a una altura física o "real" (medida sobre la superficie física de la tierra, el geoide), si se conoce el valor de dicha ondulación (Instituto Geográfico Nacional, 2009). Una forma de determinar este valor es a través de los llamados modelos gravimétricos, los cuales estiman el comportamiento de la variable a lo largo de toda superficie terrestre, siendo viable la reducción entre alturas en cada uno de los puntos que se desea posicionar en el territorio. Cabe destacar que, debido a que las dos superficies de referencia (geoide y elipsoide) utilizadas para el posicionamiento global no son paralelas, y la variación entre una y otra cambia dependiendo de la zona (Buill Pozuelo y Núñez, 2005), existen diversos modelos gravimétricos de carácter regional, local o mundial que intentan solucionar esta diferencia de magnitudes entre alturas.

Puntualmente para la Geodesia el uso de modelos geoidales es algo común, en relación a las diferentes observaciones que se realizan sobre la superficie terrestre para determinar redes geodésicas, las cuales conforman un Marco de Referencia para cualquier territorio. Estos Marcos de Referencias deben reducirse de un elipsoide a un geoide de referencia (Buill Pozuelo y Núñez, 2005) y para esto, es necesario el uso de modelos que definen las ondulaciones del geoide y sus respectivas desviaciones de la vertical a lo largo de la superficie de interés (Barahona, 2016).

Es por esto que el presente trabajo de fin de Master (TFM) tiene como objetivo la determinación de modelos locales de predicción espacial de la ondulación geoidal, utilizando como insumo la red topográfica de la ciudad de Madrid (RTM), a través de técnicas geoestadísticas en software comercial y libre (ArcGIS y R-Studio). Para generar estos modelos estocásticos en cada software primero se realizará un análisis exploratorio

de datos, seguido de un análisis estructural por medio del modelado de un variograma experimental y la predicción espacial (utilizando la interpolación kriging o “krigeaje” como método de estimación). Finalmente, los modelos obtenidos se compararán entre sí y con el modelo gravimétrico oficial para el territorio español EGM08-REDNAP (Earth Gravitational Model 08-Red de Nivelación de Alta Precisión) desarrollado por el IGN, con el fin de determinar cuál de estos modelos posee mejor precisión y se ajusta al fenómeno de la ondulación geoidal descrito por la ciudad de Madrid.

1.1 OBJETIVOS

1.1.1 OBJETIVOS GENERALES

- Modelar la ondulación geoidal de la Red Topográfica de la ciudad de Madrid (RTM), utilizando un modelo geoestadístico (Kriging) mediante software libre y comercial, para poder estimar con precisión el comportamiento de la variable de la ondulación del geoide en Madrid.

1.1.2 OBJETIVOS ESPECÍFICOS

- Modelar a través del software ArcGIS y R-Studio (<https://rstudio.com/>), un modelo local de ondulación geoidal para Madrid, utilizando técnicas geoestadísticas para generar una superficie de estimación a partir de la red topográfica de la ciudad de Madrid (RTM).
- Determinar la calidad de los modelos locales generados, por medio de una validación cruzada, comparando sus precisiones con el modelo IGN EGM08-REDNAP, el cual es de carácter nacional y oficial para España.
- Analizar la estabilidad de cada modelo resultante a lo largo de la zona de estudio, aplicando el índice de autocorrelación espacial de Moran a los residuales de la variable, para verificar la naturaleza de la superficie de estimación de los modelos.

1.2 JUSTIFICACIÓN

En la actualidad, en diferentes ciudades y países a nivel mundial se desarrollan constantemente proyectos de ingeniería o de planificación urbana donde se diseña y construye infraestructuras como edificios, puentes, vías, entre otros. Dichos proyectos en particular, y en general, en cualquier trabajo propio de las geociencias o ciencias de la tierra, es necesario conocer puntos georreferenciados sobre la superficie de la tierra que consideran la influencia de su campo gravitatorio, es decir el valor de sus respectivas alturas de carácter físicas (altura ortométrica). Para determinar estos puntos se utiliza diversos métodos, siendo el más común y actual el posicionamiento por GNSS, mismo que obtiene las 3 componentes espaciales (X, Y, Z) que describe a un punto en la superficie. Este proceso se caracteriza por sus bajos costos, rapidez y altas precisiones, mismas que están referidas a un sistema de uso global.

No obstante, el uso de técnicas GNSS a pesar de brindar a usuarios las características ya mencionadas, también posee ciertas limitaciones. Con esta técnica se determina la altura en referencia al elipsoide (altura geométrica) y no de la superficie topográfica hacia el geoide (altura física). Esta última no puede ser medida por técnicas GNSS, pero sí por técnicas topográficas como la nivelación geométrica.

La nivelación geométrica, a diferencia del posicionamiento GNSS, se caracteriza por ser una técnica costosa, tardía, con una logística compleja y de carácter regional (depende de su origen de referencia), pero necesaria si se desea conocer la altura de un punto sobre el nivel medio del mar. En otras palabras, si un usuario desea conocer las componentes $[\lambda, \varphi, H]$ de un punto sobre la superficie para aplicaciones como las descritas, necesitará de estos dos procesos geodésicos – topográficos. Lo conveniente sería poder determinar ambos componentes con una sola técnica, reduciendo así costos y tiempos operativos en campo. Por lo tanto, en referencia a lo dicho, se usará el posicionamiento GNSS por ser la técnica más eficiente, pero con una corrección o reducción de su componente vertical después de ser observado. Esto se realiza a través de la Ondulación Geoidal, misma que permite estimar una altura ortométrica en función a una altura elipsoidal de origen GNSS.

A la presente fecha existen diversos modelos gravimétricos mundiales y locales como el EGM08 (Earth Gravitational Model 2008) y el EGM08-REDNAP (modelo ajustado para España). Estos modelos tienen la finalidad de transformar las alturas mencionadas a través del cálculo de la ondulación geoidal para diferentes regiones (particulares o mundiales). Por ende, estos modelos poseen precisiones que dependerán de cómo, para qué y para quien fueron generados.

Para este TFM, como se ha mencionado, se utilizarán técnicas geoestadísticas para modelar la ondulación geoidal de la ciudad de Madrid. Esta, al ser una ciudad “cosmopolita” en constante cambio por sus diferentes dinámicas de expansión urbanística, económica y social, necesitará de infraestructura que le permita satisfacer dichas dinámicas. Estas edificaciones, al ser construidas por diversos proyectos de ingeniería, necesitarán de coordenadas conocidas con sus respectivas alturas corregidas y para ello es indispensable un modelo de ondulación geoidal local de precisión.

Para realizar la estimación de dicho modelo se utilizó como información base la Red Topográfica Municipal de la ciudad, la cual posee la altura elipsoidal y altura nivelada sobre el nivel medio del mar a través de las técnicas ya mencionadas. Finalmente, el presente documento realizará dicho modelo de estimación utilizando software libre y comercial con el fin de contrastar las ventajas y debilidades que cada uno puede ofrecer a las tecnologías de la información geográfica para desarrollar proyectos a nivel académico o laboral.

CAPÍTULO II. FUNDAMENTOS TEÓRICOS

2.1 SISTEMAS GEODÉSICOS DE REFERENCIA

La Geodesia es la ciencia que estudia la forma y dimensiones de la Tierra, su campo de gravedad y su variación temporal (Instituto Geográfico Nacional, 2009). Esta ciencia tiene el objetivo, entre otros, de determinar la posición de un objeto sobre la superficie terrestre mediante coordenadas (latitud, longitud, altura). Para lograr esto, dichas coordenadas deben estar referidas a un origen, ejes que posean dirección y sentido en referencia a superficies de revolución (figuras geométricas) (Furones, 2011). Para esto, la Geodesia ha desarrollado los denominados Sistemas de Referencia (Moirano, 2000).

2.1.1 SISTEMA DE REFERENCIA

Los Sistemas de Referencia son un conjunto de teorías y constantes que interactúan entre sí, con el fin de localizar cualquier punto en el espacio con respecto a las dimensiones descritas por superficies de referencia (comúnmente esferas o elipses). Estas están definidas por localización y orientación respecto de un eje medio, el cual coincide con el eje de rotación y el centro de masas terrestre (Hernández, S/F). En otras palabras, los sistemas de referencia deben definir un origen y orientación de manera explícita con sus respectivos modelos, hipótesis, algoritmos y constantes numéricas (Barahona, 2016).

En la actualidad, para representar la superficie de la tierra se utiliza sistemas de origen geocéntricos, mismos que son descritos a través de modelos elipsoidales de revolución (Moritz, 1984), los cuales serán detallados más adelante.

2.1.2 MARCO DE REFERENCIA

El Marco de Referencia no es más que la materialización física de un sistema de referencia en la superficie terrestre. Esta materialización se define a través de puntos fiduciales de alta precisión sobre la superficie, con coordenadas tridimensionales dadas para una época fija y sus variaciones en el tiempo (velocidades) (Drewes, 2014)

El marco no existiría sin un sistema de referencia y viceversa, generando una dependencia y retroalimentación mutua de estos dos aspectos. Esta relación define a lo que se conoce como Datum Geodésico, pues este conecta los parámetros medidos por un sistema de referencia (tamaño y orientación de un elipsoide), con los parámetros observados en puntos fundamentales en la superficie (Drewes, 2014). En términos más simples, Fernández (2012) define al datum como el punto tangente entre el elipsoide y el geoide donde estos coinciden, el cual se denomina como punto “fundamental”.

2.2 SUPERFICIES DE REFERENCIA

2.2.1 GEOIDE

De manera simple, se puede decir que el geoide es la forma real de la tierra. El Instituto Geográfico Nacional español, lo define como una superficie de nivel equipotencial del campo gravitatorio terrestre de cota cero, la cual coincide con la prolongación debajo de la superficie continental, de los mares y océanos sin el efecto perturbador de sus mareas (Instituto Geográfico Nacional, 2009).

En otras palabras, según Zakatov, 1997, el Geoide puede entenderse como la superficie que coincide con la superficie del agua en reposo de los océanos, misma que se extiende bajo los continentes de tal manera que la dirección de las líneas verticales (dirección de la plomada) son tangentes a esta superficie en todos sus puntos. Esta superficie depende de las masas en el interior de la Tierra, las cuales son desconocidas. Por lo tanto, el definir al geoide dentro de la Geodesia utilizando soluciones matemáticas, se vuelve indeterminable.

2.2.2 ELIPSOIDE

Por la irregularidad que presenta el geoide y la dificultad para definirlo, se ha visto la necesidad de emplear al elipsoide como la figura matemática simple que mejor se ajusta a la forma de la tierra. Bajo esta premisa, el centro geométrico de la figura debe coincidir con el centro de masas del planeta, al igual que sus planos ecuatoriales. El elipsoide está definido por una superficie de referencia, que se genera a partir de la rotación alrededor de su eje menor (superficie de revolución) y por los parámetros geométricos intrínsecos de la figura como lo son semiejes, achatamiento y excentricidades (Zakatov, 1997).

Cabe destacar que existe una relación entre el elipsoide y el geoide, denominada ondulación geoidal, la misma que se detalla más adelante.

2.3 SISTEMA DE COORDENADAS

Un sistema de coordenadas es usado para localizar un punto de interés sobre un espacio definido. Según Leiva (2014), para definir dicho sistema es necesario especificar:

- El tipo de coordenadas (rectilíneo, curvilíneo, plano, espacial)
- La ubicación del origen
- La orientación de los ejes
- La unidad de medida

Los sistemas de coordenadas más comunes usados para la geodesia son:

- Coordenadas cartesianas o tridimensionales $[X, Y, Z]$
- Coordenadas geográficas (elipsoidales) $[\varphi, \lambda, h_e]$
- Coordenadas planas $[x, y]$

2.3.1 COORDENADAS CARTESIANAS

Este sistema se caracteriza por estar referido al centro de gravedad de la tierra (geocentro), descrito por los ejes coordenados (X, Y, Z) (Fernández, 2012), los cuales están orientados de la siguiente forma: Eje X, sobre el ecuador y en dirección del meridiano central Greenwich. El eje Z, es coincidente con el eje de rotación de la tierra. El eje Y, es perpendicular a los otros dos ejes y coincide también en el plano ecuatorial de la tierra.

En la Figura 1, se observa el motivo por el cual este sistema de coordenadas es el más común para definir un punto en el espacio: debido a que los ejes X, Y pueden girar alrededor del eje Z por el ángulo γ , simulando la rotación terrestre y definiendo un sistema diestro, el cual gira en sentido antihorario (positivo). Así, este sistema diestro permite ubicar al punto P por un vector de posición donde x_P , y_P , z_P (componente del vector) están referidos al geocentro O (Seeber, 2003). Es importante mencionar, que estas coordenadas al ser geocéntricas y poseer ejes ortogonales entre sí, son las más utilizadas para representar coordenadas obtenidas del posicionamiento satelital GNSS y es útil para cambios entre sistemas geodésicos de referencia (Drewes, 2014).

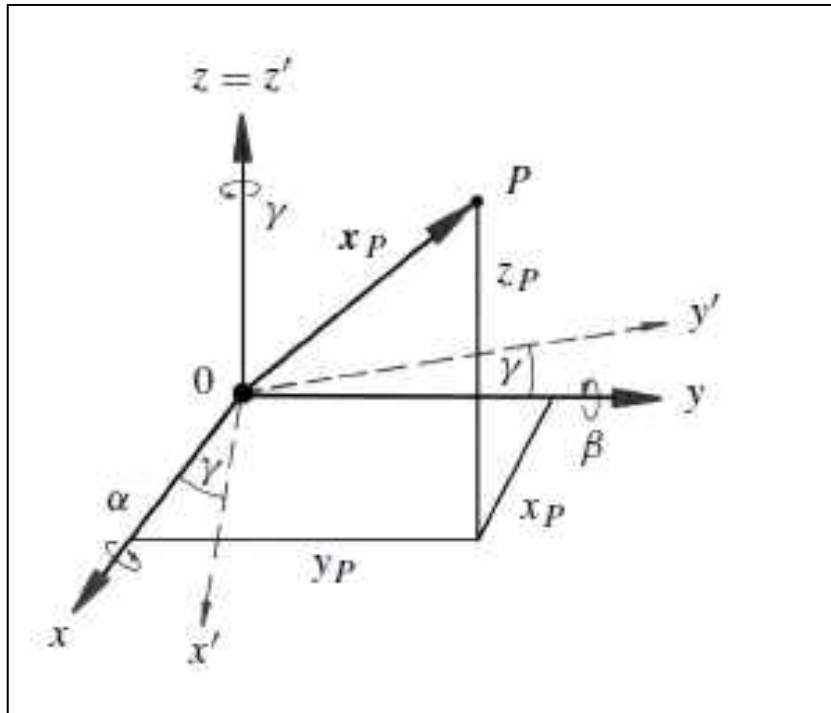


Figura 1: Sistema de coordenadas cartesianas

Fuente: (Seeber, 2003)

2.3.2 COORDENADAS GEODÉSICAS

Las coordenadas geodésicas posicionan un punto en la superficie terrestre a través de medidas angulares (latitud y longitud) referidas a un elipsoide de origen geocéntrico, como se observa en la Figura 2. Las componentes que conforman este sistema son:

- Latitud (ϕ) Angulo formado entre el plano ecuatorial con la proyección de la Normal al elipsoide que pasa por el punto P (Figura 2).
- Longitud (λ): Angulo formado entre el meridiano que contiene a P con el meridiano de Greenwich (Figura 2).
- Altura Elipsoidal (h_e): Es la distancia entre el elipsoide y el punto P , medido sobre la Normal (Figura 2).

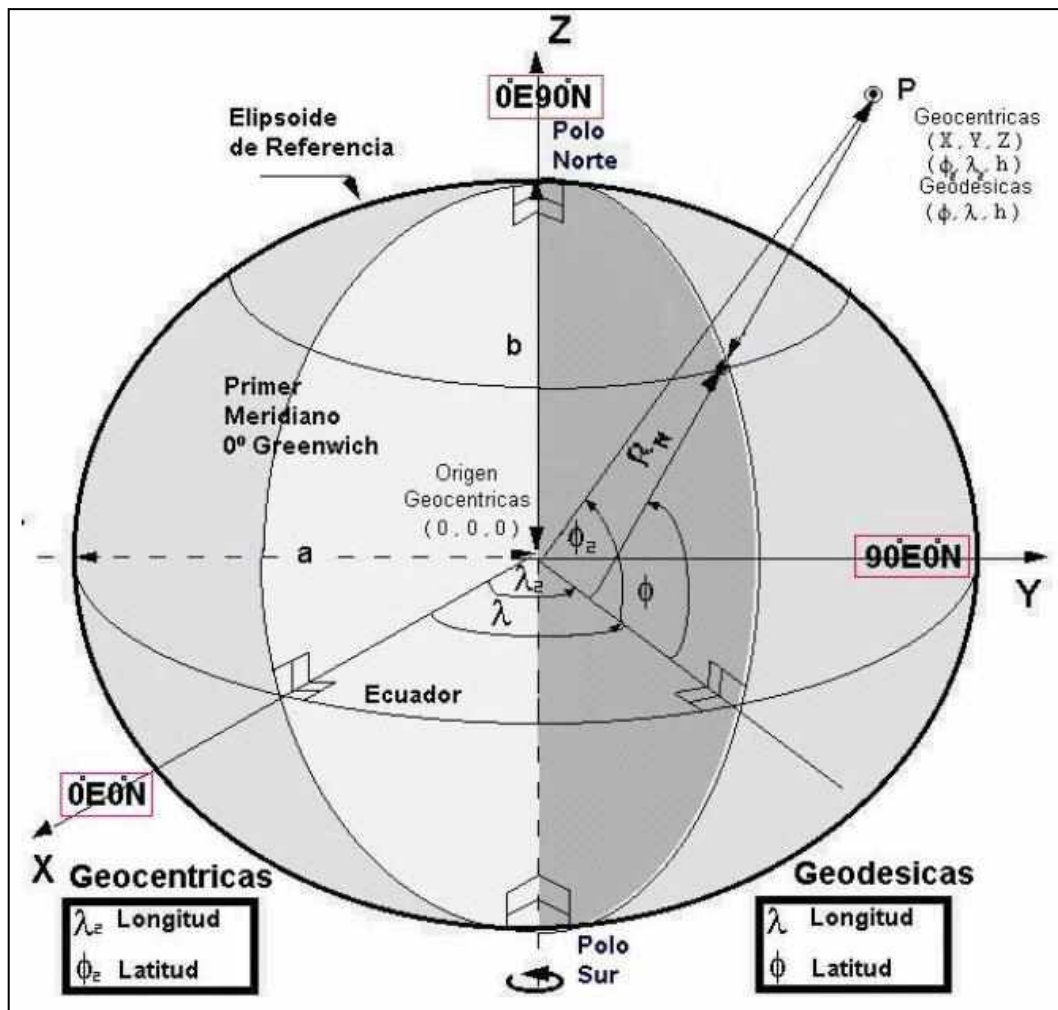


Figura 2: Sistema de coordenadas geodésicas y cartesianas

Fuente: (Fernández, 2012)

2.3.3 COORDENADAS PLANAS (UTM)

Los sistemas de coordenadas planas proyectan la superficie del elipsoide (superficies de revolución 3D) sobre una superficie plana 2D mediante transformación y reglas matemáticas o geométricas. Esta transformación entre planos siempre genera distorsiones a las dimensiones angulares, distancias o áreas. Por ello es necesario buscar un sistema que disminuya este error en función a la necesidad de cada usuario (Drewes, 2014).

Mundialmente y para el siguiente trabajo se utilizó el sistema de coordenadas UTM (Universal Transversa de Mercator), el cual es una proyección trasversa conforme (sin distorsión angular). Esta proyección es representada a través del desarrollo de un cilindro trasverso, el cual es tangente a lo largo del meridiano central de Greenwich. El problema de utilizar este tipo de proyección, es que a medida que un punto se separa del meridiano de tangencia, aumentaran las distorsiones mencionadas (Figura 3) (Hernández, S/F).

Por lo tanto, para evitar estos errores, el sistema de coordenadas UTM ha optado por un artificio, el cual consiste en subdividir la superficie de proyección en 60 husos iguales de 6° de longitud de ancho, obteniendo así 60 zonas consideradas cada una como un desarrollo cilíndrico conforme trasverso (Hernández, S/F). Cada zona a la cual llamaremos zonas UTM, estará referida a un meridiano central con el fin de reducir la distorsión en los

límites de cada zona. España se encuentra en el hemisferio norte comprendida entre las zonas 29, 30 y 31 para la península, mientras que las islas Canarias en las zonas 27 y 28.

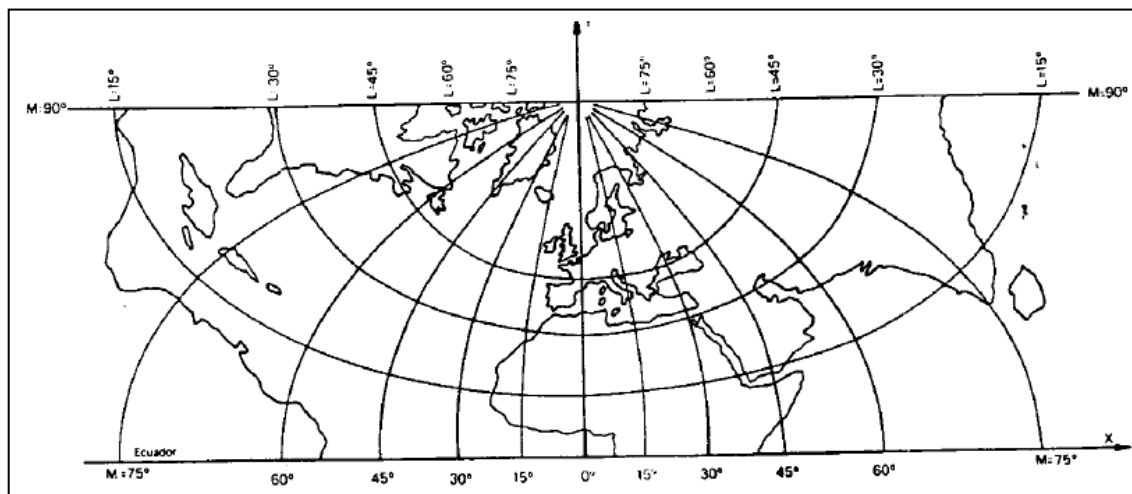


Figura 3: Distorsiones de proyección cilíndrica transversa
Fuente: (Hernández, S/F).

2.4 ALTURAS DE REFERENCIA

Como se mencionó en apartados anteriores, para poder definir un punto sobre la superficie terrestre, es necesario conocer las 3 componente $[\varphi, \lambda, h_e]$ que describen su posición. No obstante, la componente vertical (altura sobre la superficie terrestre), no solo está definida por un sistema de referencia geocéntrico como sucede con su componte horizontal (φ, λ) sino también, por la influencia de los campos gravitatorios que ejercen fuerzas sobre la superficie terrestre (Leick, 2004). En otras palabras, el valor de la altura dependerá de dos magnitudes, siendo una de ellas geométrica, la cual es obtenida a partir de diferentes técnicas topográficas y geodésicas (nivelación y posicionamiento GNSS, respectivamente). Mientras que la ortométrica, será de carácter físico y para ello deberá recibir una corrección gravimétrica (Sánchez, 2005; Leiva, 2014).

Bajo este criterio, existe diferentes tipos de alturas en función a:

- Como se obtuvo la observación en la superficie terrestre, y
- La referencia vertical (superficies equipotenciales) a la cual fue medida

Por ende, si se realizan mediciones geométricas, se obtendrá alturas nombradas de la misma manera (niveladas y elipsoidales), pero, si a estos valores se le incluye mediciones de gravedad, se obtendrá alturas físicas (ortométricas y normales) (Barahona, 2016).

2.4.1 ALTURAS ELIPSOIDALES

Las alturas elipsoidales (h_e) representan la distancia entre la superficie terrestre y el elipsoide de referencia a lo largo de la normal al mismo (ver figura 2). Esta altura es determinada a través de técnicas de navegación y posicionamiento satelital GNSS, mismas que se encuentran referidas a un sistema de coordenadas cartesianas (Barahona, 2016).

La altura elipsoidal es de carácter netamente geométrico ya que no posee valores de gravedad, pero puede tener valores similares en puntos situados en diferentes niveles. Sin embargo, también puede variar sobre una misma superficie de nivel (Freitas y Blitzkow, 1999).

2.4.2 ALTURAS ORTOMÉTRICAS

Las alturas ortométricas (H) representan la distancia entre la superficie terrestre y el geoide, a lo largo de la dirección de la gravedad (plomada). No obstante, en apartados anteriores se definieron dos conceptos de geoide: el primero, según Zakatov en 1997, considera a este como una superficie aproximada al nivel medio del mar (n.m.m) que se extiende por debajo de los continentes. En base a este concepto tradicional, el valor de H será parecido al valor de la cota medida sobre el n.m.m, mismo que está referido a un punto de la línea de costa lo más próximo al geoide (Torge, 2001).

Esta cota, es medida a través de técnicas topográficas, como la nivelación geométrica, la cual calcula el desnivel (dn) entre dos puntos A y B (figura 4). El problema de la nivelación es que considera al segmento A'B' paralelo al segmento AB para distancias cortas a lo largo de la superficie topográfica (no se considera la curvatura de la tierra a largas distancias). Por lo tanto, la diferencia entre las alturas niveladas dn no será igual a la diferencia entre alturas ortométricas dH , debido al no paralelismo de las superficies de nivel (Barahona, 2016), como se ilustra en la Figura 5.

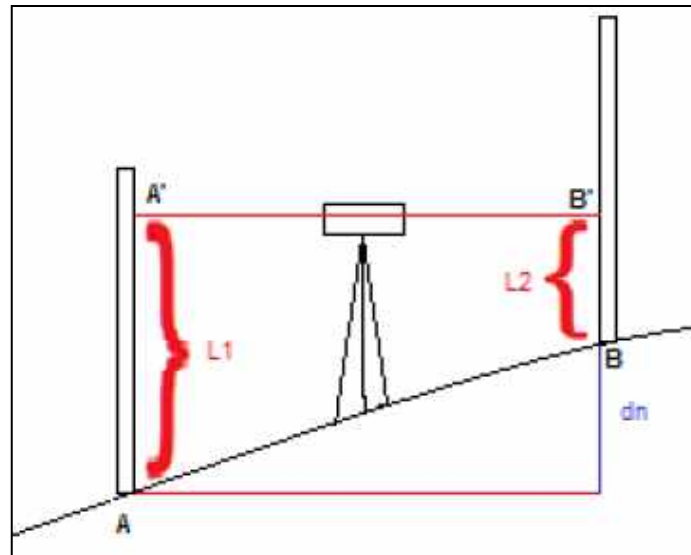


Figura 4: Nivelación geométrica

Fuente: (Barahona, 2016)

Debido al no paralelismo entre superficies, es fundamental considerar el segundo concepto (más actual) del geoide. Así, Leick (2004) lo define como una superficie equipotencial, es decir que cualquier altura que esté referida a él (en este caso H), estará influenciada por el campo gravitatorio terrestre. Por ende, el valor de la cota medida a través de la nivelación (altura geométrica) tendrá que ser corregida vinculándose con su potencial. Este último se puede definir de manera simple como, el cálculo de la intensidad de un campo de fuerza gravitacional en un punto.

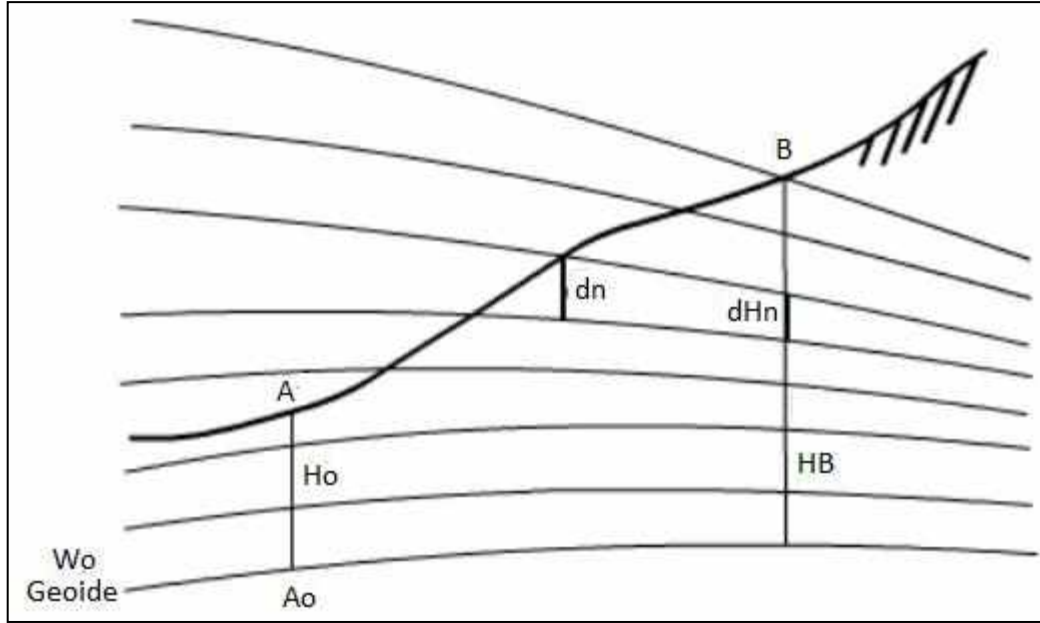


Figura 5: Diferencias entre superficies de nivel
Fuente: Adaptado de (Furones et al., 2000)

Dicha corrección parte de la premisa mencionada, en donde el incremento de nivel dn es diferente de su correspondiente dHn . Por lo tanto, según (Heiskanen y Moritz, 1985), es necesario incluir mediciones de gravedad a las diferencias de potencial (dw), de la siguiente forma:

$$-dw = g \, dn = g' \, dHn$$

En donde g es la gravedad observada simultáneamente con la nivelación y g' la gravedad media sobre la plomada de B en dHn (considerando la masa terrestre). Ahora, si se despeja dHn en función de lo planteado, se tiene:

$$dHn = \frac{g}{g'} \, dn$$

En la ecuación anterior, se observa la corrección gravimetría (g/g') que se debe aplicar en dn , para que cumpla que $dHn = dn$. Cabe mencionar, que aún es necesario conocer las diferencias de potencial entre los puntos A y B en función de la nivelación y la gravedad. Para ello volvemos a partir de su incremento dw en función de las medidas mencionadas.

$$dw = -g \, dn$$

$$WB - WA = - \int_A^B g \, dn$$

No obstante, a pesar de conocer la diferencia de potencial entre los puntos de interés, todavía no se ha referido al punto A y B al geoide de cota cero. Para ello se calcula la diferencia de potencial entre el geoide (Wo) y el punto A (WA), a lo que se conocerá como numero geopotencial y está dado por la expresión:

$$\int_0^A g \, dn = WA - Wo = C$$

C, será el número geopotencial de A, si se considera a A_0 (figura 5) como la intersección del geoide con la línea de la plomada que pasa por dicho punto. Según Furones et al. (2000) y Sánchez (2005) el número geopotencial será constante para todos los puntos de una superficie de nivel, por lo que puede considerarse como una medida física de la altura, aunque no tenga dimensiones de longitud. Por lo tanto, esta superficie considera la curvatura de la tierra y la influencia de sus campos gravitatorios, permitiéndole convertirse universalmente en el origen del datum vertical.

Finalmente, al vincular los criterios geométricos (nivelación) con los físicos (potencial), es posible calcular la altura ortométrica en función de C, mediante la expresión

$$H = \frac{C}{g'}$$

El problema con dicho cálculo radica al no conocer el valor de g' , ya que es imperativo conocer la gravedad en función de la densidad terrestre desde el punto de interés hasta el geoide a lo largo de la línea de plomada. Por lo tanto, para poder medir g' , se requiere de modelos teóricos sobre la distribución de densidad de las masas terrestres. Dicho de otra forma, la altura ortométrica dependerá de las hipótesis utilizadas en el modelamiento de la densidad terrestre, siendo las más comunes; las hipótesis de Helmert, Vignal, Baranov y Aire Libre (Free Air) (Blitzkow et al., S/F).

2.5 ONDULACIÓN GEOIDAL

Como se mencionó en apartados anteriores, la altura elipsoidal es obtenida por una solución GNSS, es decir una magnitud netamente geométrica. Sin embargo, para proyectos de ingeniería es necesario utilizar alturas relacionadas con el campo de gravedad del planeta, es decir ortométricas (Seeber, 2003).

Estas dos alturas están relacionadas entre sí, a través de la ondulación geoidal (N). Misma que, se puede definir como la diferencia entre el elipsoide y el geoide (Furones et al., 2000), como se representa en la Figura 6, y se expresa matemáticamente, por la diferencia entre la altura elipsoidal (h_e) y la altura ortométrica (H).

$$h_e = H + N$$

A través de esta relación, la Geodesia pretende calcular alturas ortométricas a través de técnicas GNSS, debido a que esta altura H, es obtenida por técnicas topográficas costosas y que toma demasiado tiempo (nivelación geométrica) (Seeber, 2003). No obstante, para cumplir con este objetivo es necesario conocer con exactitud el valor de N a lo largo de toda la superficie terrestre, considerando magnitudes físicas y geométricas como lo describe la ecuación anterior. Por lo tanto, N se deriva de modelos gravimétricos que necesitan conocer de manera precisa al geoide y sus campos de gravedad. Sin embargo, en la práctica no es así, debido a que la información actual del geoide y sus campos son insuficientes para modelar toda la superficie (Seeber, 2003; Leick, 2004). Por lo tanto, algunos autores como Zhang (2000) recomiendan realizar modelos locales para mejorar precisiones, a través de algoritmos matemáticos de interpolación entre puntos GNSS conocidos y sus alturas niveladas (por su aproximación al geoide), considerando que,

mientras mejor sea la cobertura de puntos de control, el método ofrecerá mejores resultados.

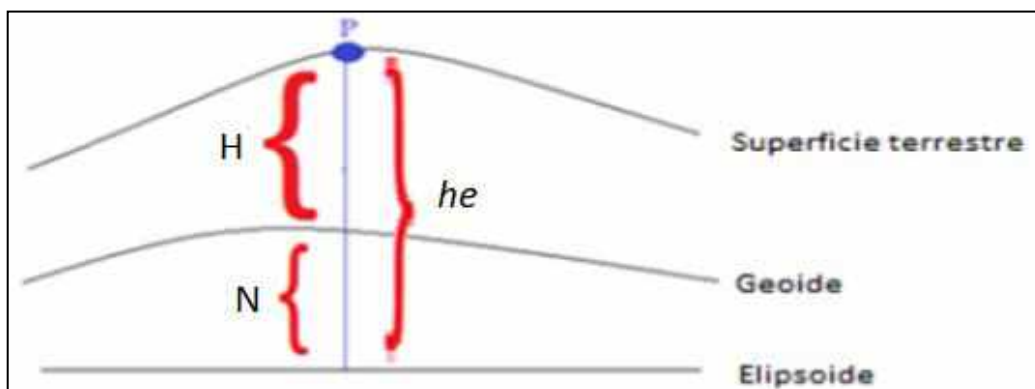


Figura 6: Alturas de referencia.

Fuente: Adaptado de (Barahona, 2016).

Es importante destacar que en la actualidad existen varios modelos de la ondulación geoidal que pueden alcanzar presiones centimétricas o menores dependiendo la zona. Todos ellos han sido creados para satisfacer distintas necesidades para múltiples usuarios y aplicaciones de ingeniería, construcción, catastro, planificación, entre otros. Estos modelos pueden ser de carácter local o global dependiendo de quién y para que fueron creados. Dicho esto, en el presente documento se hablará brevemente del modelo de ondulación geoidal oficial regido a lo largo del territorio español.

2.5.1 MODELO DE ONDULACIÓN GEOIDAL DE ESPAÑA (EGM08-REDNAP)

El Instituto Geográfico Nacional (IGN) español en cumplimiento del Real Decreto 1071/2007, definió al sistema de referencia ETRS89 y REGCAN95 como el nuevo marco geodésico oficial para el territorio español. Por lo cual, el Instituto tuvo la misión de compilar y homogenizar toda la información geodésica, topográfica y cartográfica a dicho marco (Instituto Geográfico Nacional, 2009).

En base a lo anterior, el marco de referencia horizontal actual para el territorio español se encuentra compuesto por la red geodésica de estaciones GNSS permanentes (ERGNSS) y la Red Geodésica Nacional por Técnicas Espaciales (REGENTE), las cuales proporcionan un marco homogéneo y preciso en el nuevo sistema *European Terrestrial Reference System 89* (ETRS89). Asimismo, en referencia al Artículo 4 del Real Decreto, se determina que "el Sistema de Referencia Altimétrico tomará como referencia el nivel medio del mar Mediterráneo en Alicante para la península y las respectivas referencias mareográficas...". El IGN ha integrado esta altura a sus redes geodésicas a través de modelos de ondulación geoidal, debido a que las mismas tienen alturas elipsoidales (h_e) sobre el ETRS89 bien determinadas (Instituto Geográfico Nacional, 2009).

En España, el primer modelo de geoide gravimétrico con una buena precisión fue el IBERGEO95 publicado en 1995 (referido al marco anterior, ED50) (Instituto Geográfico Nacional, 2009). No obstante, desde entonces han ido apareciendo nuevos modelos más precisos, como el EGM08, que implementa correcciones topográficas e incluye más datos de anomalías de gravedad que otros modelos predecesores.

Por ende, el IGN en su afán de poseer un sistema geodésico actualizado y homogéneo, realizó varias pruebas comparando puntos GNSS y nivelación de precisión con múltiples modelos disponibles, llegando a la conclusión que el EGM08 era el que mejor se adaptaba al marco de referencia vertical en Península y Baleares. A partir de ello, se ajustó la superficie de ondulaciones dada por un geoide gravimétrico (EGM2008) a observaciones realizadas sobre la REDNAP, con el fin de obtener un modelo resultante en el sistema de referencia materializado por el datum vertical en España (Instituto Geográfico Nacional, 2009).

En base a lo citado a lo largo del apartado 2.4, todos los modelos resultantes que consideren que dHn es casi paralelo a dn , serán aproximados. Sin embargo, como ya se mencionó antes, desde el punto de vista práctico, este criterio será suficiente para convertir alturas elipsoidales en ortométricas.

CAPITULO III. METODOLOGÍA

3.1 INFORMACIÓN Y DATOS DE PARTIDA

Sobre la base de lo planteado en los capítulos anteriores, el presente trabajo optó por tomar como fuente de información los datos proporcionados por el Ayuntamiento de Madrid y su Red Topográfica Municipal (RTM), al considerar que sirve como marco uniforme y preciso para la coordinación de todos los trabajos topográficos y productos cartográficos que se realizan a lo largo del territorio del municipio de Madrid. Si bien en la actualidad, esta Red se encuentra en un proceso de revisión y densificación en conformidad con los nuevos desarrollos urbanísticos, pero al manejar información de acceso público permite otorgar la validez necesaria al planteamiento conceptual que se presenta en este Trabajo Fin de Máster, cuya utilidad toma como premisa la última actualización que efectuó el ayuntamiento a sus datos en 2019 (Ayuntamiento de Madrid, 2019).

Con ello descrito, es necesario mencionar que la Red topográfica posee un total de 3.976 vértices materializados a lo largo de la ciudad de Madrid (excepto en el Monte de El Pardo y el Castillo de Viñuelas), referidos al sistema de referencia ETRS89, proyección UTM 30N. Estos vértices fueron densificados por posicionamiento GNSS y poligonales de precisión apoyados en la Infraestructura Geodésica Nacional (REGENTE) y en la Red de Nivelación de Alta Precisión (REDNAP) (Ayuntamiento de Madrid, 2019). La información de la RTM es proporcionada al público de manera libre en dos formatos, como fichero digital vectorial (en formato shapefile) y como fichero de texto con extensión CSV (del inglés *comma-separated values*) con la siguiente estructura:

Tabla 1: Estructura de la base de datos de la RTM.
Fuente: Municipio de Madrid.

| Campo | Tipo | Descripción |
|--------------|-------------|---|
| FID | Object Id | Código secuencial |
| Shape | Geometry | Tipo de geometría |
| Vertice | Text | Nombre de vértice |
| Hoja_MTN | Text | Hoja topográfica en la que se encuentra |
| Situacion | Text | Descripción del emplazamiento vértice |
| Señal | Text | Tipo de señal |
| Referencia | Text | Elemento de referencia y distancia |
| Referenc_1 | Text | Elemento de referencia y distancia |
| Referenc_2 | Text | Elemento de referencia y distancia |
| Referenc_3 | Text | Elemento de referencia y distancia |
| X_ED50 | Double | Coordenada X en ED50 |
| Y_ED50 | Double | Coordenada Y en ED50 |
| X_ETRS89 | Double | Coordenada X en ETRS89 |
| Y_ETRS89 | Double | Coordenada Y en ETRS89 |
| Z_Ortometr | Double | Altura nivelada |
| Altura | Double | Altura elipsoidal |
| AzmVert1Vi | Text | Azimut visual 1 |
| Nombre1Vis | Text | Número vértice visual 1 |
| AzmVert2Vi | Text | Azimut visual 2 |
| Nombre2Vis | Text | Número vértice visual 2 |
| AzmVert3Vi | Text | Azimut visual 3 |
| Nombre3Vis | Text | Número vértice visual 3 |
| AzmVert4Vi | Text | Azimut visual 4 |
| Nombre4Vis | Text | Número vértice visual 4 |
| Año | Long | Año de revisión |
| Fecha | Date | Fecha de revisión |

3.2 OBTENCIÓN DE LA ONDULACIÓN GEOIDAL DE LA RED TOPOGRÁFICA MUNICIPAL DE MADRID (RTM)

Como se puede observar en la Tabla 1, no existe el atributo de la ondulación geoidal para cada punto o vértice de la RTM; por lo tanto, es necesario calcular dicho valor. Para efectuar aquello, es necesario realizar una diferencia entre la altura elipsoidal con la nivelada (considerada como ortométrica), teniendo en cuenta que estas alturas están vinculadas mediante la ondulación geoidal (N) a través de una relación geométrica (Barahona, 2016).

Por lo tanto, a la tabla de atributos de base RTM se le agregó un campo con el cálculo de N en función de la diferencia de los valores de las otras alturas, es decir $[Altura - Z_{Ortometr}]$ (ver figura 7). Este paso es fundamental pues permite llevar a cabo el análisis, depuración y exploración de datos de la variable de interés, previo a iniciar un modelado con Kriging.

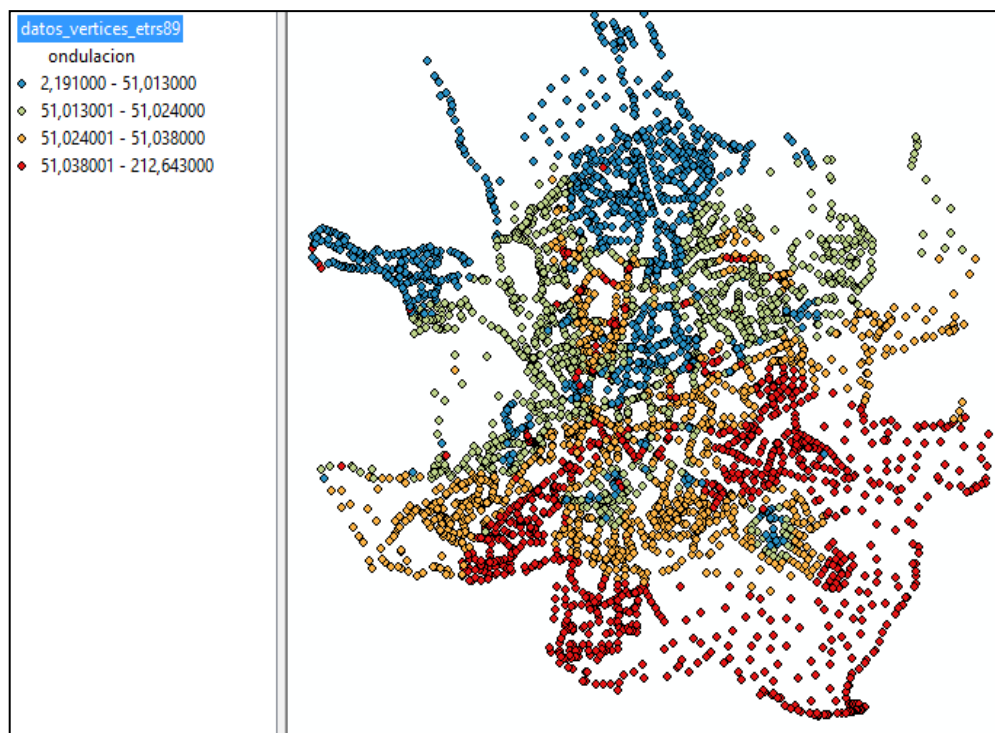


Figura 7: Ondulación geoidal RTM por cuartiles, sin depurar.

3.3 ANÁLISIS Y DEPURACIÓN DE LOS VALORES DE ONDULACIÓN GEOIDAL DE LA RTM.

Con lo anterior detallado, es imperativo efectuar una revisión y limpieza de los datos obtenidos, asegurando así la calidad del producto generado en el presente trabajo. Para esto se utilizó el Sistema de Información Geográfica ArcGIS. Esta herramienta permite el análisis y depuración de los datos de ondulación geoidal, mediante la visualización de la distribución espacial de la variable. Por ende, se utilizaron las herramientas de estadística descriptiva que posee este SIG para identificar patrones de comportamiento, datos atípicos y clusters.

Como muestra en las Figuras 7 y 8, se puede observar de forma respectiva la presencia de datos atípicos en los extremos de la tendencia. El 50% de los datos se concentran entre los intervalos 51.013 y 51.038 metros de N (Q1 y Q3, respectivamente), mientras que sus valores mínimos y máximos son de 2 y 212 metros respectivamente.

Para limpiar esta base de datos, se eliminaron los vértices que estaban fuera de rango del 95% de su diagrama de caja, más una desviación estándar (valores 51 ± 5 m, figura 9). Estos vértices se caracterizaban por poseer una distribución espacial aleatoria en relación a su entorno, como se muestra en la Figura 10; de esta forma se eliminaron 37 puntos de 3976, siendo estos definidos como errores groseros dentro de base de la RTM.

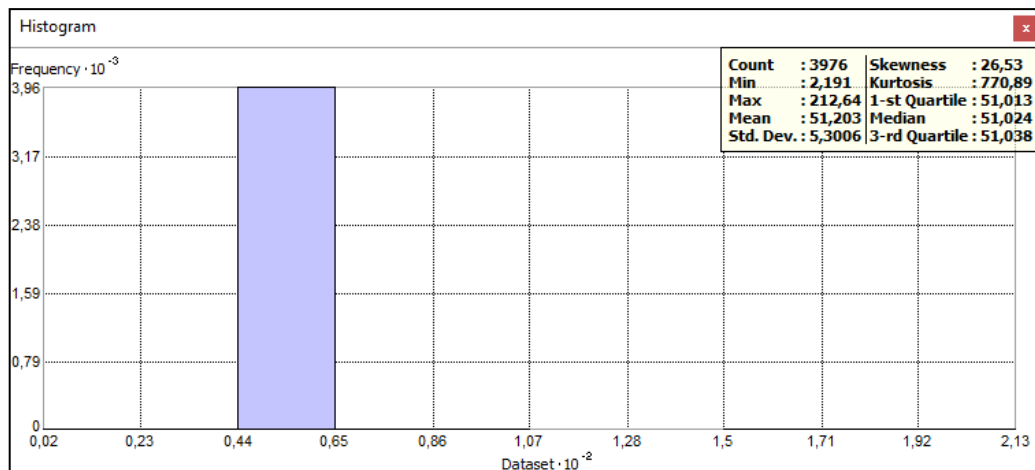


Figura 8: Histograma y descripción de N con datos crudos, en ArcGIS.

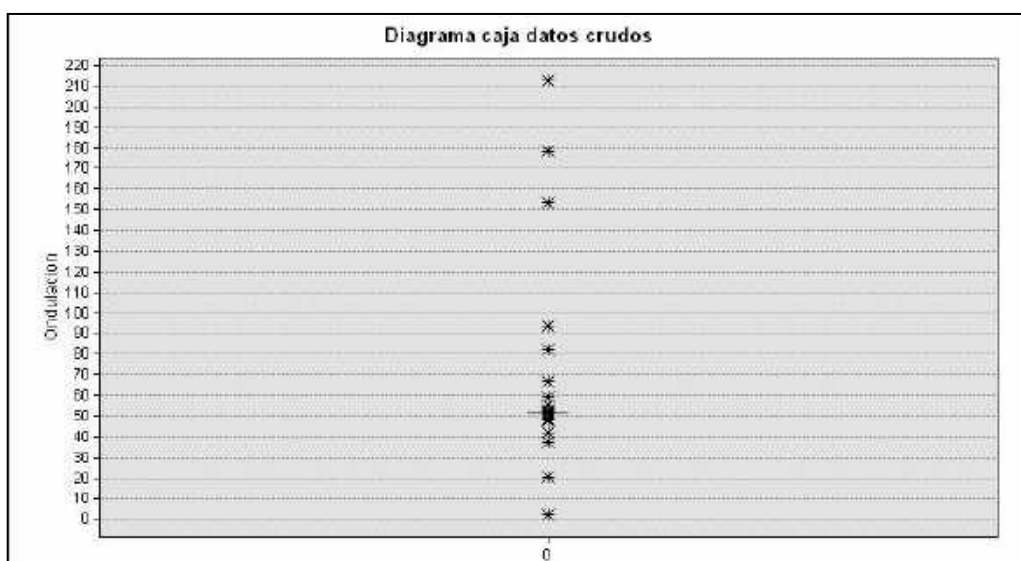


Figura 9: Diagrama de caja y bigote datos crudos.

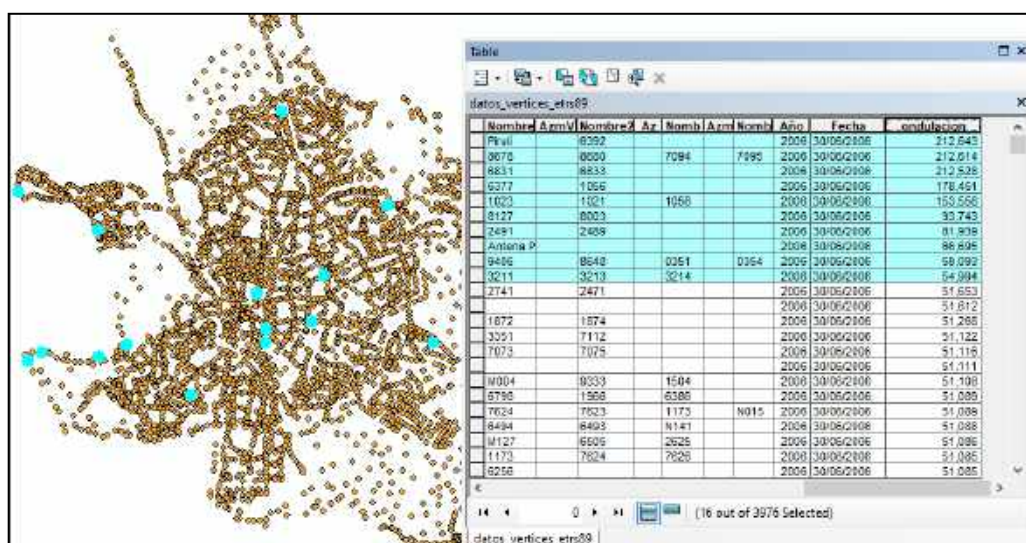


Figura 10: Selección de datos atípicos de una desviación estándar positiva.

Una vez depurados los datos se obtuvo una distribución más homogénea que varía en milímetros, incluso en aquellos valores que se encontraban fuera del 95% de la muestra,

cómo se observa en la Figura 11, en el que ambos extremos no superan el 1.5 mm de diferencia del límite de su segmento o “bigote” respectivo. A su vez, el diagrama de caja con datos depurados posee bigotes más equidistantes entre sí con respecto a los valores de tendencia central, lo que indica una cohesión y comportamiento estable de los datos, por lo tanto, no se consideró necesario eliminar más. Cabe destacar que los datos atípicos son valores inusualmente grandes o pequeños comparados con el resto de la muestra, así que es necesario tener en cuenta qué rangos son los adecuados para eliminarse, ya que esta acción puede afectar la interpretación y manejo de los mismos.

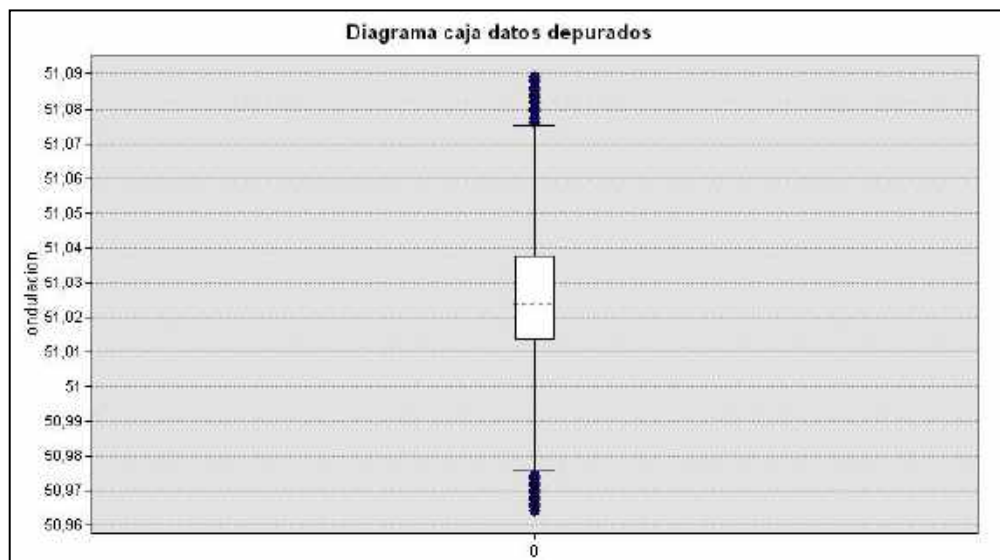


Figura 11: Diagrama de caja y bigote de datos depurados

3.4 MÉTODOS DE INTERPOLACIÓN (KRIGING) Y ESTADÍSTICA ESPACIAL PARA ESTIMAR LA ONDULACION GEOIDAL EN LA SUPERFICIE DEL MUNICIPIO DE MADRID.

En este apartado se explican los principales métodos, técnicas y criterios de estimación geoestadísticos utilizados en el presente TFM, para modelar la ondulación geoidal del término municipal de Madrid, a través de un SIG (ArcGIS) y en un programa estadístico (R-Studio).

3.4.1. MÉTODOS DE INTERPOLACIÓN

La interpolación espacial es un proceso que permite estimar o calcular el valor de una variable en el espacio, conociendo los valores de otras posiciones (Bosque Sendra, 2000). Estos métodos se clasifican en dos grandes categorías: determinísticos y estocásticos (probabilísticos).

3.4.1.1. Métodos determinísticos

Los métodos de interpolación determinísticos calculan un valor aproximado, usando funciones matemáticas que describen las propiedades o relaciones de las variables involucradas (Goovaerts, 1997). Los métodos de interpolación determinísticos más usados son la interpolación por el inverso de la distancia (IDW), triangulación (polígonos Thissen o Voronoi), regresiones polinómicas (trend surface analysis, funciones spline) (Chica Olmo y Luque Espinar, 2002), interpolación del vecino natural, entre otros (Paredes et al., 2012).

3.4.1.2 métodos estocásticos

Los métodos de interpolación probabilísticos, a diferencia de los determinísticos, estiman una gran cantidad de información utilizando la variación y correlación estadística de las muestras disponibles (Maune et al., 2001). Los métodos de interpolación probabilísticos más comunes son los estadísticos espaciales (geoestadísticos), de donde se deriva el método Kriging (Paredes et al., 2012).

3.4.2 ESTADÍSTICA ESPACIAL

Según Giraldo (2007), la estadística espacial es un conjunto de metodologías que analizan datos que corresponden a variables aleatorias en diversos sitios (esto es característico de puntos en el espacio o agregaciones espaciales) de una región. Dicho de otra forma, la estadística espacial analiza las relaciones de un proceso estocástico de una variable en un espacio euclídeo (espacio geométrico que está representado por una superficie de 1, 2 o 3 dimensiones).

3.4.2.1 Geoestadística

La geoestadística es una rama de la estadística que analiza fenómenos espaciales, siendo su objetivo principal la estimación, predicción y simulación de dichos fenómenos (Giraldo, 2007). Por otra parte, Emery (2013), describe a la geoestadística como el estudio de fenómenos regionalizados, es decir, fenómenos que se desarrollan en un espacio geográfico y presentan cierta continuidad. Estos fenómenos muestran dos características complementarias: por una parte, tienen una cierta “continuidad” espacial (posee zonas de valores altos o bajos); y, por otra, varían irregularmente, lo que provoca que no se puedan representar de manera simple.

Debido a la extrema complejidad de representar una variable regionalizada, Emery (2013) recomienda no modelar estas variables por medio de funciones determinísticas, sino por estimaciones probabilísticas. Para ello, se justifica el uso de la geoestadística, la cual trabaja en dos etapas. La primera es el análisis estructural, que describe la correlación entre puntos observados en el espacio (muestras del fenómeno), mientras que la segunda predice en lugares no muestreados, utilizando la técnica denominada Kriging (Oliver y Webster, 2010). Kriging es un proceso que calcula la media ponderada de las observaciones, asignando pesos a los valores muestrales en función de la estructura espacial de correlación establecida en la primera etapa (Giraldo, 2007).

3.4.2.2 Geoestadística y la teoría de las variables regionalizadas

Alfaro (2007), al igual que Emery (2013), definen la geoestadística como “la aplicación de la teoría de las variables regionalizadas a la estimación de fenómenos aleatorios, conociendo que una variable regionalizada es una función que representa la variación en el espacio de una cierta magnitud asociada a un fenómeno natural”.

Dicho de una manera más formal, la variable regionalizada es aquella que se puede medir en el espacio y posee una estructura de correlación (Giraldo, 2007). Es decir, un proceso estocástico con dominio en un espacio euclídeo d -dimensional R^d , donde $Z(x)$ representa una variable medida en un punto y considerando en este caso $d=2$.

$$\{Z(x): x \in D \subset R^d\}$$

Las observaciones en campo de las variables regionalizadas suelen estar próximas entre sí y varían en función a su posición. Es decir, están correlacionadas, lo que permite describirlas por la estimación de su covariancia.

$$C(x_1, x_2) = E[\{Z(x_1) - \mu(x_1)\}\{Z(x_2) - \mu(x_2)\}]$$

$\mu(x_n)$ representan las medias de $Z(x)$
E: denota el valor esperado

Cabe mencionar, que la geoestadística necesita de una nube densa de puntos, que abarque el área más grande posible para representar al fenómeno de interés. Sin embargo, la variable regionalizada suele poseer muestras dispersas a lo largo de la zona de estudio, con una sola una observación por punto (Emery, 2013).

En razón a esto, si se analiza la ecuación anterior, solo se tendrá un valor de $Z(x)$ en cada punto, siendo imposible calcular las medias de la covariación (solución no viable). Por lo tanto, para solucionar este problema se debe utilizar supuestos de estacionariedad que se explican a continuación (Oliver y Webster, 2010).

3.4.2.3 Estacionariedad

Como se mencionó en el apartado anterior, la existencia de una observación por punto impide el cálculo de la covarianza de la variable aleatoria. Para ello, la estacionariedad pretende inferir estadísticamente la repetición de observaciones (es decir más de un valor para $Z(x)$), por medio de una repetición en el espacio. Esta última afirmación según Emery (2013), parte de la hipótesis que se supone que “los valores que se encuentran en las diferentes regiones del campo, presentan las mismas características y, por ende, pueden considerarse como distintas observaciones del mismo proceso aleatorio”. En otras palabras, la estacionariedad postula que el valor de la variable es invariante con respecto a su posición, como se ilustra en la Figura 12, en donde el valor $Z(x)$ es similar a $Z(x + h)$ para cualquier distancia h , permitiendo obtener un valor medio (círculo rojo) y constante de estos valores.

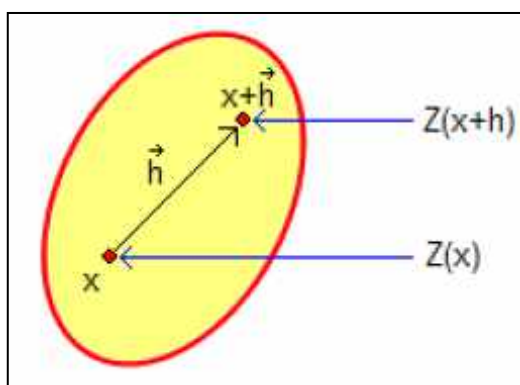


Figura 12: Dos puntos a una distancia vectorial h
Fuente: (Alfaro Sironvalle, 2007)

Por lo tanto, la teoría de la estacionariedad modificará a la covarianza en razón a las siguientes tres condiciones (Giraldo, 2007; Oliver y Webster, 2010). La primera, si su media $\mu = E[Z(x)]$, es constante para todo x , entonces $\mu(x_1)$ y $\mu(x_2)$ pueden ser remplazados por μ , la cual debe ser estimada. La segunda, si $x_1 = x_2$ entonces la correlación espacial definida por la ecuación de la covarianza, ahora es definida por la

varianza $\sigma^2 = E[\{Z(x) - \mu\}^2]$, que se supone es finita y como la media, constante en todo el conjunto de puntos. Finalmente, la tercera si $x_1 \neq x_2$ su varianza depende de h y no de sus posiciones absolutas, para cualquier par de puntos x_i, x_j separados por la distancia $h = x_i - x_j$, de manera que:

$$\begin{aligned} C(x_i, x_j) &= E[\{Z(x_i) - \mu\}\{Z(x_j) - \mu\}] = \\ &= E[\{Z(x)\}\{Z(x+h) - \mu\}] = \\ &\equiv C(h) \end{aligned}$$

La ecuación anterior, denota que la covarianza depende directamente de los valores de $Z(x)$ en función de la distancia h . Estas tres condiciones de la covarianza según Giraldo (2007) se conocen como estacionariedad de segundo orden.

3.4.2.3.1 Estacionariedad Intrínseca o Débil

Existen algunos fenómenos físicos reales en los que la varianza no es finita (procesos no estacionarios), en donde su media no es constante y debido a ello, no existe un valor de μ para poder determinar la ecuación $C(h)$ y su covarianza respectiva (Oliver y Webster, 2010). A este problema se lo conoce como estacionariedad intrínseca, que según Clark (1979), se soluciona con la hipótesis que; si a la variable $Z(x)$ se le agrega incrementos $[Z(x+h)]$ éstos serán estacionarios siempre y cuando los valores de h sean lo más pequeños posibles, volviendo a las diferencias esperadas nulas, de manera que:

$$E[Z(x+h) - Z(x)] = 0$$

Al remplazar las covarianzas por la esperanza de las diferencias al cuadrado, se obtiene:

$$var[Z(x+h) - Z(x)] = E[\{Z(x+h) - Z(x)\}^2] \equiv 2\gamma(h)$$

De esta manera la covarianza y la varianza, esta última conocida ahora como variograma, dependen solamente de h y no de las posiciones de los puntos (Oliver y Webster, 2010). Ahora bien, si el proceso $Z(x)$ es estacionario de segundo orden, el variograma y la covarianza serán equivalentes. Por otra parte, si dicho proceso es no estacionario, no habrá dicha equivalencia porque la función covarianza no existe. Sin embargo, se puede utilizar en su lugar el variograma (Giraldo, 2007).

3.4.3 ANÁLISIS EXPLORATORIO DE DATOS

En la geoestadística es imperante, así como en otros procesos estadísticos, el análisis gráfico, y el estudio de los datos y su tendencia, ya que estas herramientas permiten identificar valores extremos, datos atípicos, variabilidad, correlación y patrones en su ubicación geográfica de la variable regionalizada de interés (Giraldo, 2007).

Según Alfaro (2007), la importancia de este análisis es primordial para establecer si algunos fenómenos físicos se ajustan a la teoría geoestadística, misma que permitirá definir que procedimiento de predicción es el más conveniente para su representación.

3.4.4 ANÁLISIS DE ESTRUCTURA

3.4.4.1 Variograma experimental

Según lo visto anteriormente, la función variograma $2\gamma(h)$ constituye la herramienta fundamental de la geoestadística (Alfaro, 2007). No obstante, a la mitad de este $\gamma(h)$, se lo conoce como la estimación de la semivarianza o variograma experimental, el cual no se trata de una función como tal, sino de una serie de valores que se calculan para vectores de posición h , mismos que pretenden definir las propiedades de correlación espacial de un proceso $Z(x)$ (Emery, 2013).

$$\gamma(h) = \frac{\sum [Z(x+h) - Z(x)]^2}{2n}$$

Donde según se había definido, $Z(x)$ es el valor de la variable en una posición X y $Z(x+h)$ en un punto separado del anterior por una distancia h ; n , es el número de parejas que se encuentran separadas por dicha distancia, como se representa en la Figura 13.

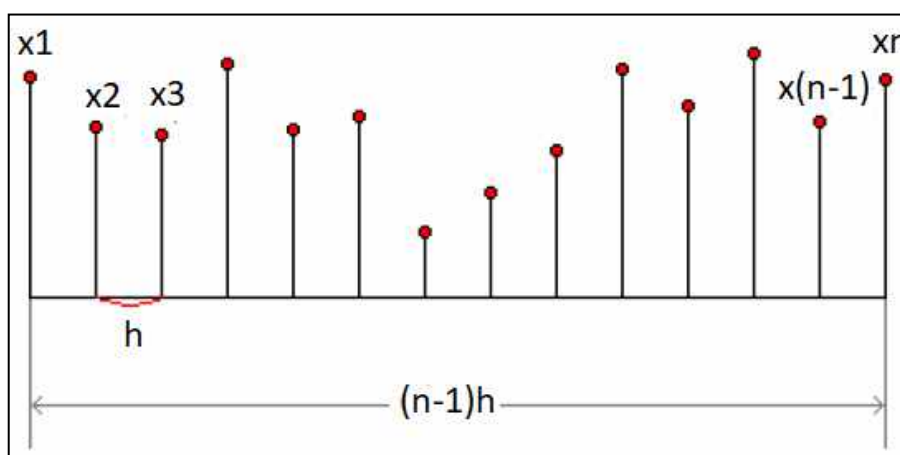


Figura 13: Disposición de muestras separadas por h

Fuente: Adaptado de (Alfaro Sironvalle, 2007)

La función semivarianza debe calcularse para varias distancias h . No obstante, en la práctica, debido a la irregularidad de las muestras y sus distancias, se toman intervalos de $\{[0, h], (h, 2h], (2h, 3h], n\}$, obteniendo un semivariograma experimental que corresponde a una distancia promedio entre pares de puntos. Obviamente el número de parejas de puntos n dentro de los intervalos no es constante (Giraldo, 2007).

Para Alfaro (2007), la interpretación del semivariograma experimental considera que “a menor distancia entre los sitios, mayor similitud o correlación espacial entre las observaciones”. En otras palabras, una variable presenta autocorrelación si los valores de h son directamente proporcionales a los valores de su semivarianza.

3.4.4.2 Covariograma

Según Giraldo (2007), la función de covarianza entre pares de puntos separados por una distancia h (ver figura 18) se calcula, empleando la fórmula clásica de la covarianza muestral:

$$C(h) = \frac{\sum [Z(x+h) * Z(x)]^2}{n} - m^2$$

Donde m representa el valor promedio en todo punto de la región de estudio y n es el número de pares de puntos que se encuentran a una distancia h . Para Oliver y Webster (2010) se debe considerar que bajo un supuesto de estacionaridad, cualquiera de las dos funciones mencionadas (covarianza o semivariograma), pueden ser usadas para definir la relación espacial de una variable. No obstante, la semivarianza no requiere estimación de otros parámetros (m). Por ende, Oliver y Webster (2010) ratifican lo dicho por Giraldo (2007), al mencionar que en la práctica es preferible usar la semivarianza, en vez de la covarianza.

3.4.4.3 Modelos teóricos de la semivarianza

Definir el variograma experimental es solo la primera aproximación del fenómeno en estudio, debido a que este es determinado solo para un cierto número de distancias, direcciones y datos observados (Emery, 2013). Dicho esto, es necesario ajustar el variograma experimental a un modelo teórico, el cual busca predecir la continuidad espacial de la variable, a través de la técnica de interpolación kriging.

Como se mencionó antes, el semivariograma calcula la autocorrelación en función de una muestra de distancias medias (h) dentro de una zona limitada, convirtiéndose en una estructura empírica o experimental. Por ello, se ajustará dicha estructura a modelos que generalicen los puntos observados en el semivariograma a cualquier distancia, como se ilustra en la Figura 14.

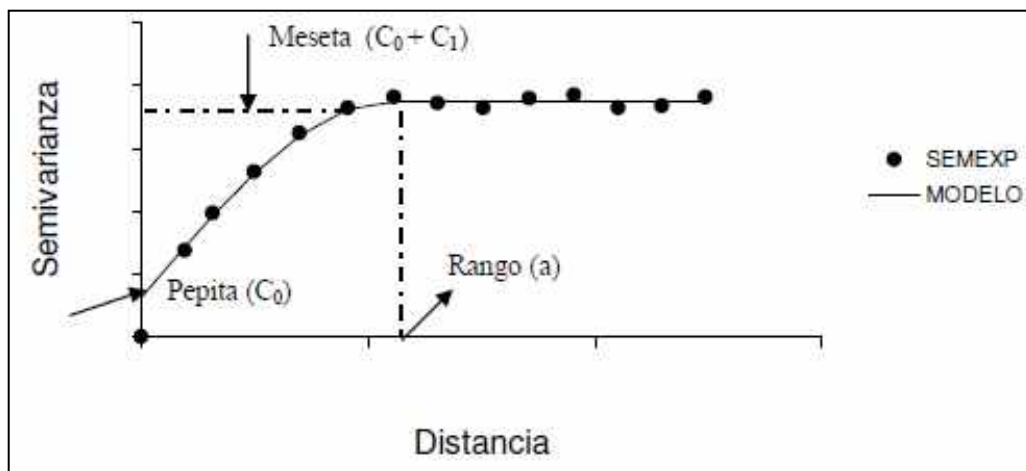


Figura 14: Estructura general de un semivariograma. SEMEXP corresponde al semivariograma experimental y MODELO al ajuste de un modelo teórico.

Fuente: (Giraldo, 2007)

Estos modelos siempre presentaran las siguientes características:

- **Efecto Pepita (Nugget effect):** Denotado por (C_0) y representa una discontinuidad puntual del variograma en el origen. Esto se puede generar por errores sistemáticos al momento de medir la variable en campo.
- **Meseta (Sill):** Es la cota superior del variograma. También puede definirse como el límite del variograma cuando la distancia se incrementa demasiado, provocando que la

meseta pueda o no tender al infinito. La meseta se denota $(C_0 + C_1)$ cuando la pepita es diferente de cero.

- **Rango (Range):** En términos prácticos corresponde a la distancia partir de la cual dos observaciones son independientes, en otras palabras, hasta qué punto la variable tiene autocorrelación espacial. Se lo representa con (a) .

Finalmente, es necesario resaltar que existen diversos modelos teóricos de semivarianza que pueden ajustarse al semivariograma experimental. En este TFM se analizó la estructura de los modelos más comunes, según Giraldo (2007) y Emery (2013) (figura 15), que se explican brevemente a continuación:

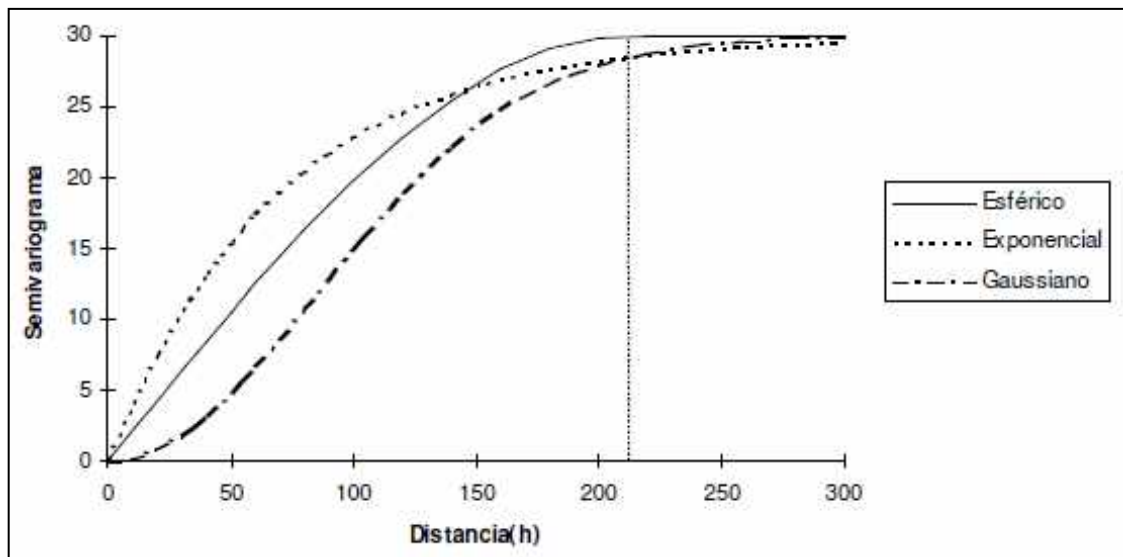


Figura 15: Comparación de los modelos teóricos exponencial, esférico y gaussiano.

Fuente: (Giraldo, 2007)

3.4.4.3.1 Modelo Exponencial

Este modelo se aplica cuando la dependencia espacial tiene un crecimiento exponencial respecto a la distancia entre las observaciones.

$$\gamma(h) = C_0 + C_1(1 - \exp\left(\frac{-3h}{a}\right))$$

3.4.4.3.2 Modelo Gaussiano

Para este modelo, la dependencia espacial disminuye con distancias que tienden al infinito. La característica principal de este modelo es su forma parabólica cerca al origen.

$$\gamma(h) = C_0 + C_1(1 - \exp\left(\frac{-h^2}{a^2}\right))$$

3.4.4.3.3. Modelo Esférico

Este modelo se caracteriza por tener un crecimiento rápido cerca al origen, pero sus incrementos extremos decrecen para distancias grandes, hasta superar el valor del rango, en donde se vuelven nulos. Su expresión matemática es:

$$\gamma(h) = Co + C1 \left(\frac{3}{2} * \frac{h}{a} - \frac{1}{2} * \left(\frac{h}{a} \right)^3 \right) \quad si \ h \leq a$$

$$\gamma(h) = Co + C1 \quad si \ h > a$$

3.4.4.3.4 Modelo efecto pepita puro

Este modelo representa la falta de correlación espacial entre las observaciones de una variable (figura 16).

$$\gamma(h) = 0 \quad si \ h = 0$$

$$\gamma(h) = Co \quad si \ h > 0$$

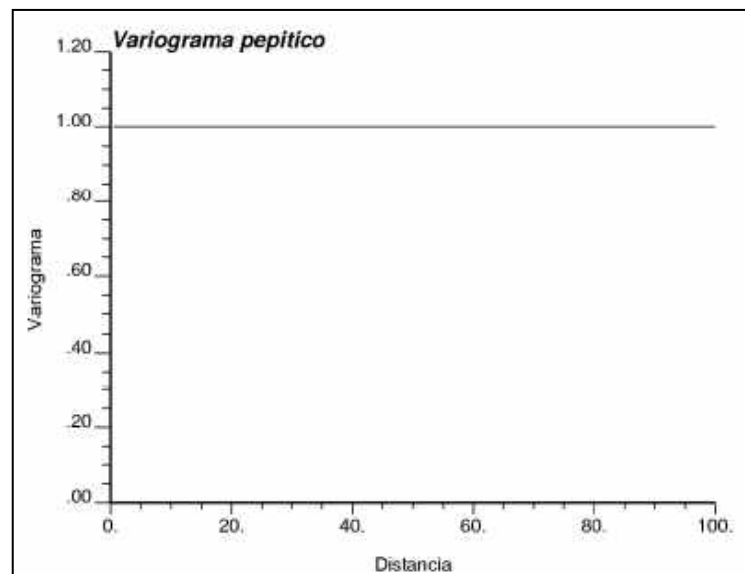


Figura 16: Modelo pepita puro.

Fuente: (Emery, 2013)

3.4.5 PREDICCIÓN ESPACIAL KRIGING

Kriging es un conjunto de métodos de predicción espacial que se fundamenta en el principio de minimizar el error medio cuadrático (RMSE, *Root Mean Square error*) de predicción (Giraldo, 2007). Además, Kriging no solo ofrece predecir valores, sino también las varianzas Kriging o también conocido como errores de precisión del modelo. Dicho de una manera más formal, Kriging puede ser considerado como un método que calcula un promediado ponderado de los valores observados de una variable aleatoria $Z(x)$, dentro de una vecindad móvil $V(x)$ (Oliver y Webster, 2010). En este sentido, Emery (2013) recomienda utilizar la vecindad móvil cuando se tiene datos “cercaños” al sitio a estimar. El objetivo es generar varias estimaciones locales a través de una grilla regular, que cubra toda la zona de estudio.

Existen varios tipos de kriging, cada uno con sus diferentes propiedades de estimación en función del tipo de variable que se desee modelar, siendo principalmente los siguientes:

3.4.5.1 Kriging Ordinario

Para este método, se considera que existe mediciones de la variable de interés Z en los puntos x_i , donde $i = 1, 2, \dots, n$. Es decir, se conoce la región de estudio definida por $Z(x_1)$, $Z(x_2)$..., $Z(x_n)$, a la cual se desea predecir $Z(x_0)$, en un punto x_0 donde no hubo medición. Bajo este supuesto, el método kriging ordinario propone, que el valor de la

variable puede predecirse como una combinación lineal de las n variables aleatorias bajo la siguiente expresión (Giraldo, 2007).

$$Z(x_0) = \sum_{i=1}^n \lambda_i Z(x_i)$$

En donde λ_i representa los pesos o ponderaciones de las variables medidas en campo. Dichos pesos se calculan minimizando el valor de la varianza, en función de la distancia entre los puntos muestreados y el punto que se desea predecir, es decir a partir del semivariograma. Esta suma debe ser igual a 1, para que la esperanza del predictor sea igual a la esperanza de la variable (Giraldo, 2007). Este método kriging es el más común para predecir todo tipo de variables y el que mejor se ajusta a múltiples procesos estacionarios o estacionarios intrínsecos (en especial el este último).

3.4.5.2 Kriging Simple

Para este método, se debe considerar que existe una variable regionalizada estacionaria con media (m) y covarianza conocidas. De manera similar a como se define en modelos lineales, este modelo será igual a la media más un error aleatorio que tiende a cero $E[\varepsilon(x_i)] = 0$, con la característica que estos errores no son independientes (Giraldo, 2007).

$$Z^*(x_0) = m + \sum_{i=1}^n \lambda_i \varepsilon(x_i)$$

Por lo tanto, a diferencia del kriging ordinario, este método no posee restricciones para las ponderaciones, es decir su suma no será igual a la esperanza de la variable.

3.4.5.3 Kriging Universal

En los casos anteriores se ha asumido que la variable regionalizada es estacionaria (a menos que se cumpla con la hipótesis intrínseca). No obstante, para aquellos casos en donde la variable no satisface estas condiciones y se caracteriza por exhibir una tendencia, Giraldo (2007) recomienda el uso del kriging universal. Para este método, se descompone a la variable $Z(x)$ como la suma de la tendencia, tratada como una función determinística $m(x)$, más una componente estocástica estacionaria cercana a cero, $E[\varepsilon(x)] = 0$ y $V[\varepsilon(x)] = \sigma^2$ (Giraldo, 2007).

$$\begin{aligned} Z(x) &= m(x) + \varepsilon(x) \\ Z(x) &= m(x) \\ m(x) &= \sum_{i=1}^p a_i f_i(x) \end{aligned}$$

En donde las funciones $f_i(x)$ son conocidas y p es el número de términos empleados para ajustar $m(x)$. Ahora, el estimador del kriging universal será el mismo usado para el kriging ordinal (es decir $Z(x_0)$), siempre y cuando se mantenga el criterio que la esperanza del predictor sea igual a la esperanza de la variable (este criterio se conoce como insesgado). Para ello $Z(x_0)$ debe cumplir la siguiente condición (Giraldo, 2007).

$$\begin{aligned}
E[Z(x_0)] &= m(x_0) \\
E\left[\sum_{i=1}^n \lambda_i Z(x_i)\right] &= m(x_0) \\
\left[\sum_{i=1}^n \lambda_i m(x_i)\right] &= m(x_0) \\
\left[\sum_{i=1}^n \lambda_i\right] \sum_{i=1}^p a_i f_i(x_i) &= \sum_{i=1}^p a_i f_i(x_0) \\
\sum_{i=1}^p a_i \left[\sum_{i=1}^n \lambda_i f_i(x_i)\right] &= \sum_{i=1}^p a_i f_i(x_0) \\
\sum_{i=1}^n \lambda_i f_i(x_i) &= \sum_{i=1}^p f_i(x_0)
\end{aligned}$$

Finalmente cabe destacar que $m(x)$ puede ser un polinomio de coordenadas, funciones trigonométricas o una función externa independiente. Por lo tanto, la dificultad que implica utilizar este método kriging, radica en poder ajustar el modelo, debido a que la función descrita por $m(x)$ provoca que el variograma experimental este sesgado con respecto al variograma teórico (Emery, 2013).

3.4.6 VALIDACIÓN KRIGING

Existen diferentes métodos para determinar la calidad de ajuste de un modelo generado por kriging con respecto a los datos observados. El más empleado es la validación cruzada, misma que excluye la observación de uno de los n puntos observados y con los $(n-1)$ valores restantes genera un modelo de predicción, para posteriormente estimar el valor de la variable en la ubicación del punto que se excluyó y compararlos (Giraldo, 2007).

Este procedimiento se debe realizar en forma secuencial con cada uno de los puntos muestrales (incluidos y excluidos dentro del semivariograma), obteniendo un conjunto de n errores de predicción o también denominado residuales. Una vez se han obtenido estos errores de predicción es común utilizar un criterio de valoración como es el Error Medio Cuadrático (RMSE, por sus siglas en inglés) con el objetivo de probar diferentes modelos de semivarianza para seleccionar aquel que optimice dicho criterio (Giraldo, 2007).

CAPITULO IV. RESULTADOS

4.1 MODELO KRIGING USANDO EL SOFTWARE ARCGIS

Como paso previo a iniciar el proceso de realizar un modelo estocástico de N, se consideró necesario dividir en dos conjuntos aleatorios los datos originales, que fueron previamente depurados. El objetivo de crear conjuntos, es usar uno de ellos para modelar la variable de estudio, mientras que el otro permite comprobar y validar la calidad de dicho modelo.

- El primer conjunto: Fue una muestra aleatoria (80% de la población) que se utilizó para modelar la estructura espacial y producir una superficie de estimación.
- El segundo conjunto: Fue otra muestra aleatoria (20% de la población) que se empleó para comparar y validar los resultados obtenidos de la superficie estimada (validación cruzada).

Una vez separados los datos, se procedió a diseñar y construir el mencionado modelo mediante técnicas geoestadísticas como el kriging. Para esto, se requiere de varias etapas fundamentadas en los criterios descritos en el capítulo anterior, cuya aplicación y resultados se muestran en los siguientes subapartados.

4.1.1 ANÁLISIS EXPLORATORIO DE DATOS

En esta primera etapa de depuración y análisis exploratorio de datos, se utilizó las herramientas *Explore data* del módulo de Geoestadística Espacial (*Geostatistical Analyst*) de ArcGIS, con el fin de identificar y eliminar datos atípicos, analizar la tendencia y comportamiento de los datos a lo largo de la superficie de estudio.

Los resultados muestran que los valores de N depurados (primer conjunto de datos separados) del municipio de Madrid presentan una tendencia centralizada bastante homogénea (media y mediana son muy similares), tal como se había reflejado y analizado en el diagrama de caja de la Figura 16. La distribución de los datos también presenta una campana de Gauss asimétrica positiva de curtosis leptocúrtica, gracias a la pequeña variación de la desviación estándar, con valores inferiores a los 2 cm, tal como se observa en la Figura 17.

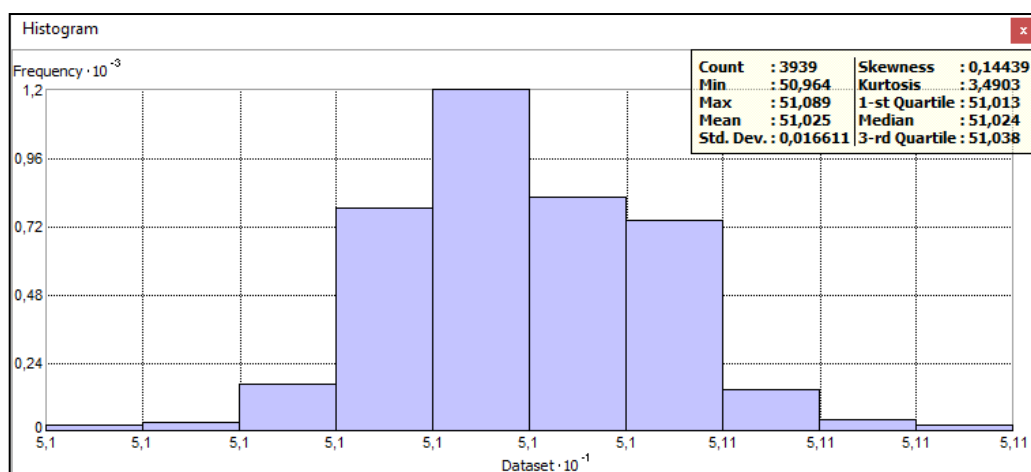


Figura 17: Histograma de ondulación geoidal con datos depurados, en ArcGIS.

Posteriormente, se procedió a analizar la distribución de los datos originales comparados con su respectiva normalización, para así determinar una posible tendencia de los

residuales del modelo a evaluar. En la Figura 18, se puede evidenciar que los valores residuales tienden a crecer en los valores extremos (subestimando en los valores altos y sobreestimando en los valores bajos); a ese fenómeno se lo conoce como heteroscedasticidad, el cual se define como la varianza no constante de los residuales de un modelo de estimación (Arce, 2001).

Sobre este punto, al considerar la posible presencia de dicho fenómeno en el presente trabajo fin de Máster, primero se tiene que modelar la ondulación geoidal, para después a través del Índice de Moran aplicado a los residuales (se detalla en el apartado 4.5), comprobar si existe significancia estadística o no de este fenómeno.

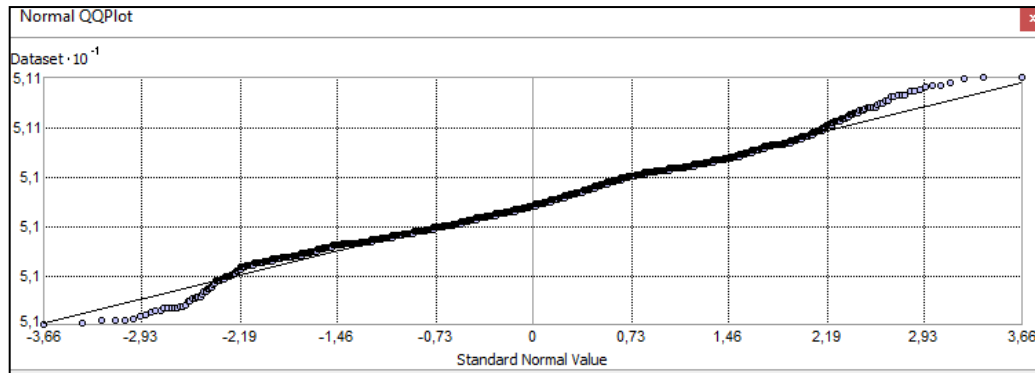


Figura 18: Gráfico de normalidad de N.

Siguiendo con los resultados del análisis exploratorio de datos con el módulo de ArcGIS, se observa que los datos al estar depurados, no presentaran zonas con valores atípicos en su retardo espacial (spatial Lag). Esto se manifiesta en el mapa de Voronoi con los valores medios de N (no es constante en la región) de la Figura 19, donde se observa una variación paulatina y continua de los valores. Incluso para ratificar este comportamiento se calculó el coeficiente de variación de los datos, obteniendo así un valor inferior al 0.1%.

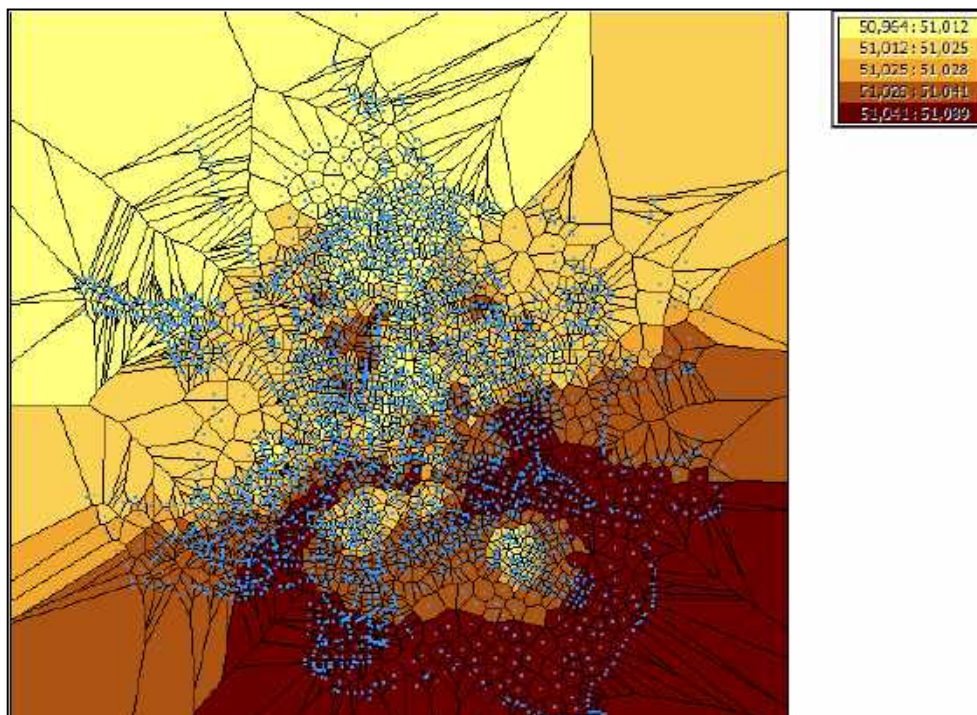


Figura 19: Mapa de Voronoi de la media de N.

Además, como parte del análisis también es necesario determinar si la variable espacialmente representada posee alguna tendencia de estacionaridad y de ser así, en que grado lo hace, considerando lo que indica Giraldo (2007) “la estacionaridad existe cuando la esperanza matemática (media) de la variable es constante para toda la región”. Al tomar como base dicha afirmación y a sabiendas que el análisis debe realizarse sobre cada componente espacial de la variable, se utilizó la herramienta de tendencia 3D del módulo *Geostatistical Analyst*. El resultado demostró que los valores de N son procesos no estacionarios (representado por funciones acotadas) de segundo orden, en donde su componente X tiene una dependencia creciente, mientras que la componente Y es decreciente, tal como lo refleja la Figura 20.

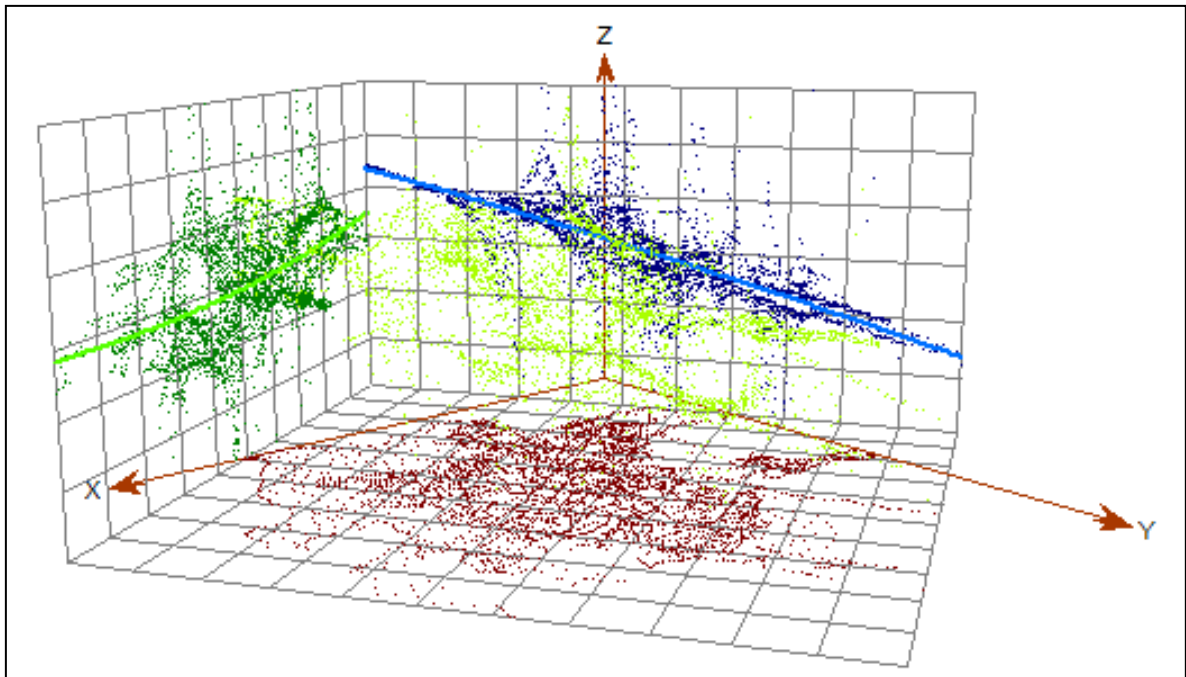


Figura 20: Estacionaridad de N en proyección 3D. Verde: Proyección XZ, Azul: Proyección YZ, Rojo: Proyección XY, Amarillo: Proyección XYZ.

4.1.2 VARIOGRAMA EXPERIMENTAL

Antes de construir un Variograma en función de la variable de estudio, se debe tener en cuenta la primera ley de la geografía expuesta por Waldo R. Tobler's (1970), quien argumenta que “todos los objetos están relacionados, pero los objetos más cercanos entre sí, están más relacionados que los distantes.” Esta ley permite definir la autocorrelación espacial, la cual es esencial para la geoestadística (Alfaro, 2007).

La autocorrelación espacial existe por la relación funcional que hay de un punto determinado con respecto a su entorno, esta relación puede ser descrita por una función semivariograma que depende de una distancia definida (ver apartados 3.4.2.3 y 3.4.4). Esta distancia (h o también conocida como Lag) es valorada de manera bidimensional ya que una variable espacial tendrá como mínimo dos dimensiones (su posición en X y su posición en Y). Por lo tanto, esta distancia será representada como un vector (con coordenadas cartesianas 2D o polares, Figura 21) (Alfaro, 2007).

Por otra parte, definir correctamente el valor de Lag o retardo espacial de una función semivariograma, es de suma importancia. Ya que este parámetro define el valor máximo que podrá alcanzar una Meseta o Sill para funciones acotadas, mas no, para funciones no acotadas (Oliver y Webster, 2010). Por lo tanto, una manera de determinar el tamaño del retraso espacial es usar la herramienta *Average Nearest Neighbor* de ArcGIS, la cual calcula una distancia promedio entre puntos y sus vecinos más cercanos. Esto ayuda a seleccionar un tamaño de retraso razonablemente bueno, ya que cada retraso tendrá al menos un par de puntos (ESRI, 2019).

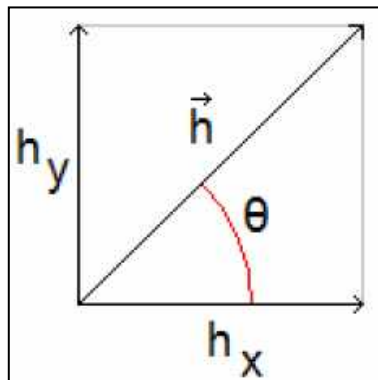


Figura 21: Componentes del vector h . θ es el ángulo de coordenadas polares

Fuente: (Alfaro Sironvalle, 2007)

$$h = (h_x, h_y) \leftarrow \text{coord. cartesianas}$$

$$h = (|h|, \theta) \leftarrow \text{coord. polares}$$

Como se mencionó antes, un Variograma se construye en función de h , pero h al ser un vector dependiente de θ que posee un dominio de 360° , puede generar 360 variogramas de forma respectiva, que describen la autocorrelación de una variable. No obstante, no todas las variables presentan cambios significativos en función de la dirección de análisis del variograma. Es decir, no es necesario analizar 360 veces un fenómeno. Según (Isaaks y Srivastava, 1989) este fenómeno de variación en función de la dirección del análisis se lo conoce como isotropía y anisotropía.

Cuando la dependencia espacial varía según la dirección del análisis, el proceso es anisotrópico, mientras que si la dependencia espacial es aproximadamente igual en las direcciones de análisis se lo entiende como un proceso isotrópico (Emery, 2013; Leiva, 2014). Para determinar si un proceso es isotrópico o anisotrópico, Alfaro (2007) recomienda tan solo analizar 4 direcciones, considerando a las demás posibilidades innecesarias. Estas son:

- $\theta = 90^\circ$, es decir, la dirección en sentido Norte - Sur
- $\theta = 0^\circ$, es decir, la dirección en sentido Este - Oeste
- $\theta = 45^\circ$, es decir, la dirección en sentido diagonal (Noreste)
- $\theta = 135^\circ$, es decir, la dirección en sentido diagonal opuesto (Suroeste)

Teniendo como base estas consideraciones teóricas, en este trabajo se analizaron diferentes variogramas en función de patrones de vecindad, con un ángulo de barrido de 45° y un Lag de 200 a 250 metros aproximadamente en las siguientes direcciones: 0° , 90° , 45° y 135° . Los resultados obtenidos se muestran en la Figura 22.

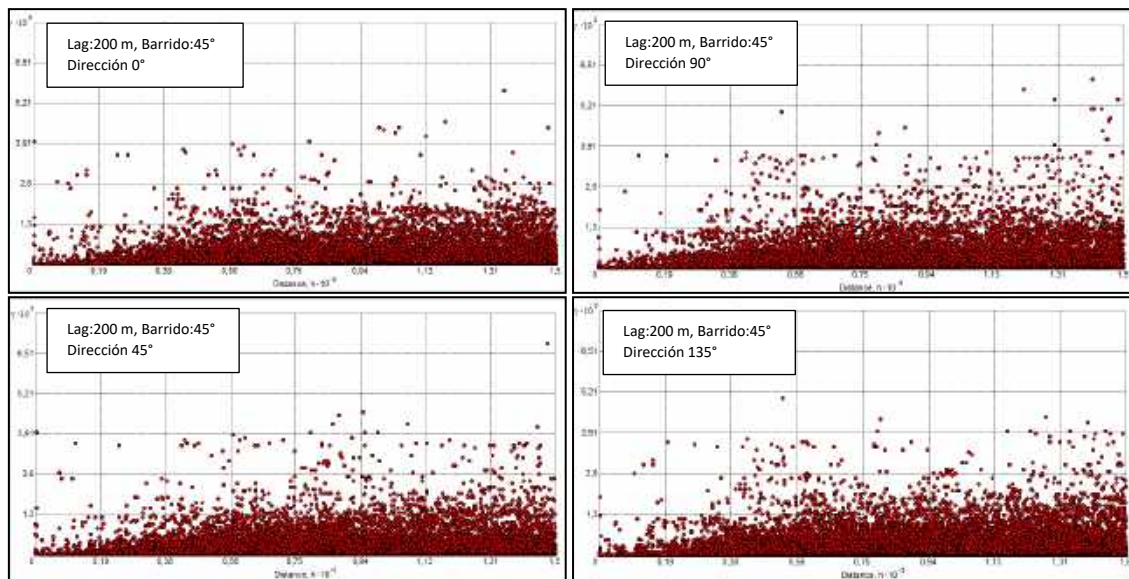


Figura 22: Variograma con Lag de 200 a 250 m con ángulo de barrido de 45° y de direcciones 0°, 90°, 45° y 135° respectivamente.

Al comparar los 4 variogramas experimentales que se generaron por cada patrón de búsqueda, se determina que existe isotropía (Variograma omnidireccional) de una función semivarianza acotada. Eso se debe a que los valores de la varianza empírica casi no varían en función a la dirección de h . En relación a esto, es perentorio comentar que existe un problema en el momento de modelar el variograma experimental usando ArcGIS, ya que no se puede visualizar la línea de tendencia que describe la semivarianza. Esto dificulta el análisis a razón de que solo se puede observar la semivarianza como una nube de dispersión, justamente como lo refleja la figura 22. De forma afortunada, dicha tendencia se volverá visible en el siguiente apartado.

4.1.3 VARIOGRAMA TEÓRICO

La finalidad de esta etapa es crear un modelo teórico (función de predicción) que se ajuste y describa el comportamiento del variograma experimental. Para ello se utilizó la herramienta *Geostatistical Wizard* del módulo de *Geostatistical Analyst* de ArcGIS, pues este permite probar diferentes modelos teóricos que describan el comportamiento de la variable en estudio.

En la Figura 23 se puede examinar la tendencia del variograma experimental representada por cruces azules, mientras que los puntos rojos representan muestras de la nube de dispersión de la Figura 22. Gracias a la mencionada tendencia del variograma se obtuvieron, como resultados, las siguientes características (para profundizar ver apartado 3.4.4.3):

- Presenta efecto pepita (Nugget Effect) aproximado de 0.00003 m^2
- El valor de alcance (Rango) aproximado es de 3000 m.
- El valor de la meseta (Sill) aproximada es de 0.00019 m^2 .
- La forma en el origen es parabólica (similar al modelo teórico Gaussiano)
- Por la forma del variograma experimental, los modelos teóricos que podrían ajustar a esta estructura son: gaussiano, circular, exponencial y cúbico.

De igual manera la Figura 23, permite observar la configuración del ajuste que tendrá el variograma teórico usando ArcGIS, representado por la línea azul, bajo los siguientes parámetros:

Tabla 2: Parámetros del ajuste del variograma teórico al experimental

| Propiedades función semivariograma teórica | |
|--|----------------------|
| Función | Gaussiana + Circular |
| Isotropía | Si |
| Efecto pepita | Si |
| Lag | 250 |
| Numero de Lag | 12* |
| Patrones de búsqueda | |
| Ancho de banda | 3 |
| Angulo de búsqueda | 45° |
| Dirección de búsqueda | Omnidireccional |

***Nota:** Numero de Lag seleccionado por defecto. Sin embargo, autores como Alfaro (2007), Giraldo (2007) y Leiva (2014) recomienda que sea 10.

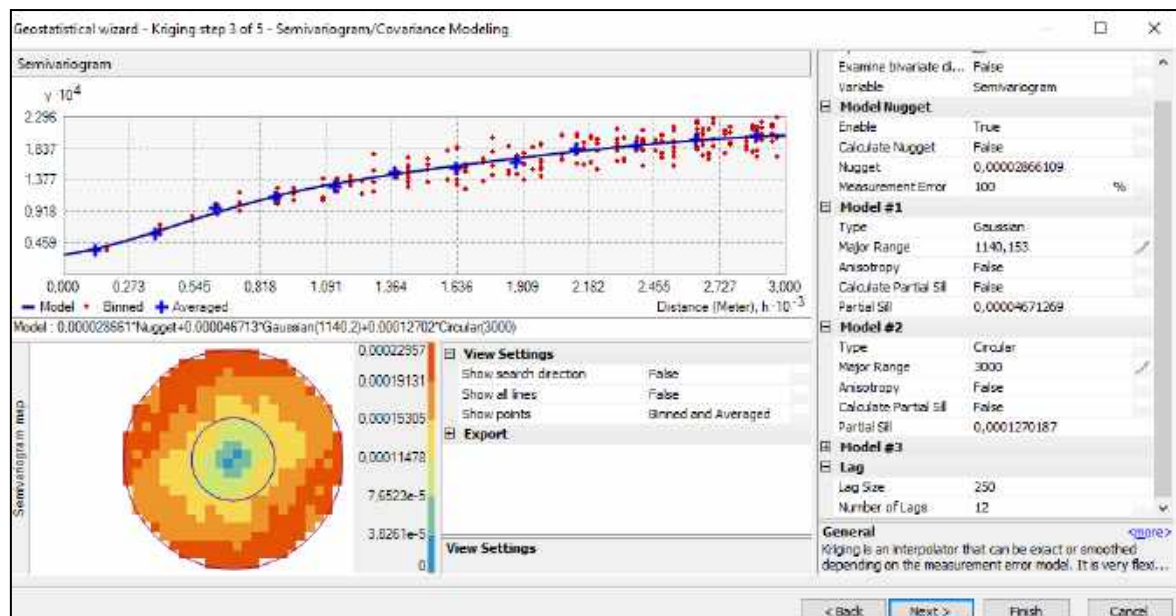


Figura 23: Modelado de variograma teórico de N en ArcGIS

Después de realizar algunas pruebas, este Trabajo Fin de Máster propone ajustar el modelo teórico con una función compuesta bajo el principio de parsimonia, para obtener una estructura básica o una combinación lineal de funciones (modelos anidados) que mejor se ajusten al variograma experimental (Emery, 2013; Desassis y Renard, 2013). En la Tabla 3, se muestra los parámetros del modelo anidado que mejor se ajustó al variograma experimental.

Tabla 3: Parámetros de Estructuras Básicas - ArcGIS

| Modelos Básicos | Parámetros | |
|-----------------|---------------|------------------------|
| | Rango (m) | Sill (m ²) |
| Gaussiano | 1140,15 | 0,00004671269 |
| Circular | 3000,00 | 0,00012701870 |
| Efecto pepita | 0,00002866109 | |

4.1.4 VALIDACIÓN CRUZADA INTERNA

Esta técnica divide los datos en dos muestras, escogiendo una para medir la precisión de la predicción del ajuste, mientras que la otra será para estimarla (Cañadas, 2013). ArcGIS crea de manera automática una validación cruzada, pero lo realiza con los mismos datos con los que construyó el modelo de predicción. Por lo tanto, a este proceso “interno” hecho por el software se la denominará validación cruzada interna.

Para este proceso es necesario definir previamente una vecindad de búsqueda (matriz de búsqueda), la cual seleccionará la muestra (vecinos cercanos) para determinar la calidad del ajuste, gracias a la diferencia del dato observado con su homólogo estimado (Leiva, 2014). Como se mencionó antes, el software ArcGIS lo hace de manera autónoma en función a la cantidad y distribución espacial de los datos de entrada.

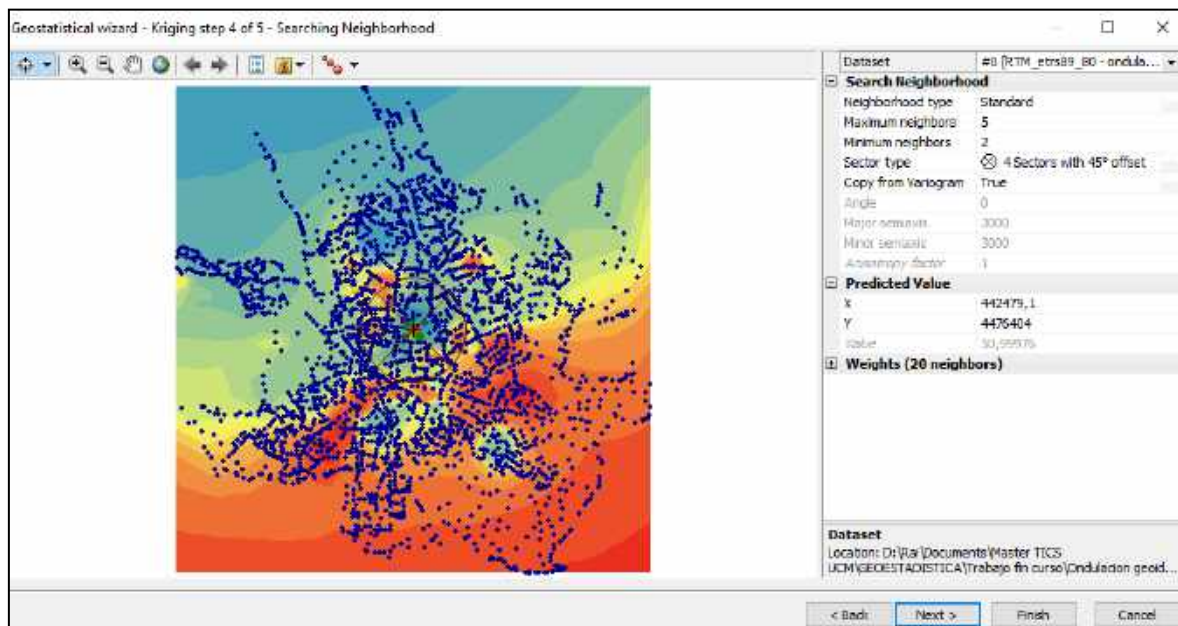


Figura 24: Matriz de búsqueda definida por ArcGIS

Es necesario resaltar que la matriz de búsqueda desempeña un papel primordial, pues es el método de interpolación que usará el software para representar la superficie de estimación (en función del semivariograma teórico), tal como se verá más adelante. Por ello, en la Figura 24 se señala los parámetros definidos por el software para crear la matriz de búsqueda, mismos que se enlistan a continuación:

- Mínimo número de muestras en la vecindad: 2
- Máximo número de muestras en la vecindad: 5
- Definición de sectores: Sí (para desagregación de las muestras)
- Número de sectores: 4 (cuadrante) con un offset de 45°
- Máxima distancia de búsqueda: 3000 m en función del variograma experimental.

Finalmente, la Figura 25 muestra los valores de la validación cruzada interna. Para esto se analizaron 3151 datos, obteniendo un RMSE observado de 5.5 mm así como un estandarizado muy cercano a 1 (el modelo interpreta al 83% de los datos). Sin embargo, se observa que la distribución de los errores no es homogénea en los extremos del modelo (presenta heteroscedasticidad, como se evaluó en la figura 18). Este problema demanda realizar un análisis del índice de Moran aplicado a los residuales de la validación cruzada interna como externa. Esto se detallará en el apartado 4.5.

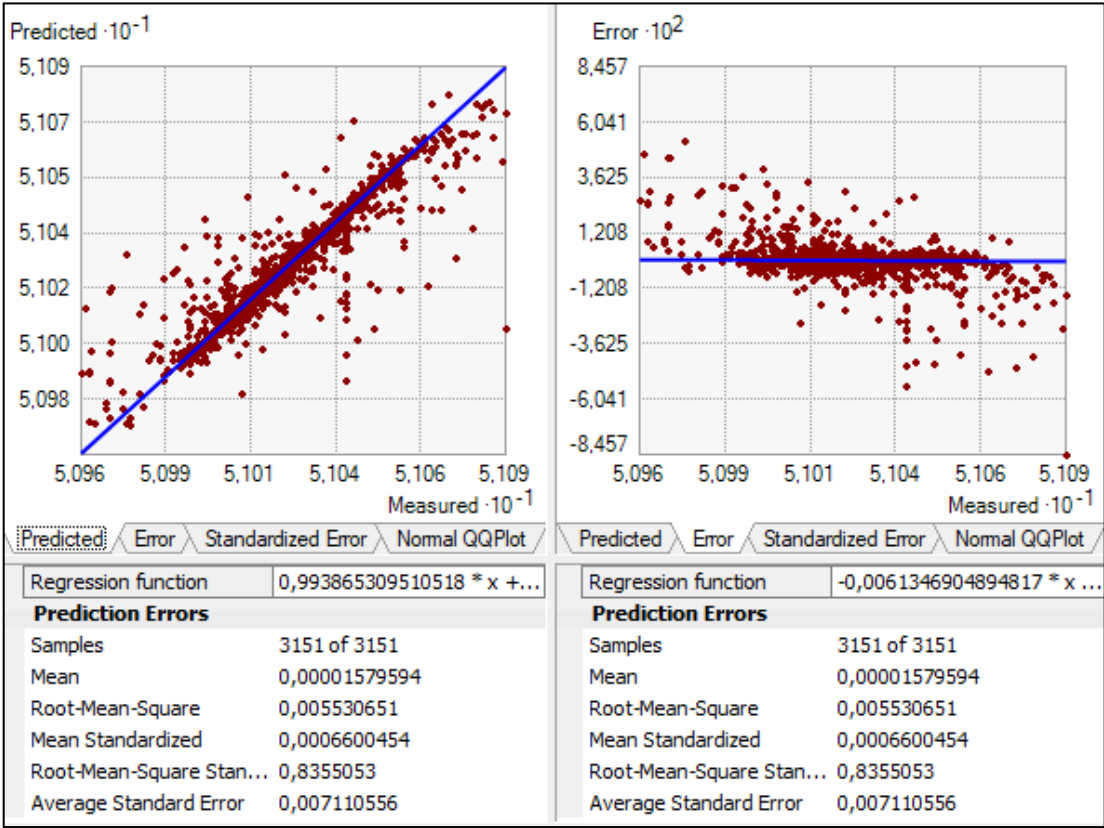


Figura 25: Validación cruzada interna por ArcGIS

4.1.5 SUPERFICIE DE PREDICCIÓN

La superficie de predicción espacial de la ondulación geoidal para el municipio de Madrid se realizó a través de un modelo estocástico de estimación Kriging ordinario puntual, exportado sobre una malla regular de puntos con intervalo de 50 m (resolución espacial) en formato IMG. Dicha resolución fue determinada por 1/4 de intervalo de la densidad espacial media que posee la RTM (aproximadamente 1 vértice cada 200 m), además el coeficiente de variación de la variable de estudio es inferior al 0.5%, lo que hace incensario aumentar la resolución espacial.

La estimación espacial se obtuvo con la función kriging de la herramienta *Geostatistical Wizard* del módulo *Geostatistical Analyst* de ArcGIS, con los parámetros definidos en el apartado 4.1.3, y con un método de interpolación por vecindad detallado en el apartado 4.14. Como resultado final se consiguió una función Kriging en formato raster, en donde sus niveles digitales representan los valores de predicción de N para el municipio de Madrid (figura 26 y Anexo 4), además de su respectiva desviación estándar para cada pixel (figura 27 y Anexo 5).

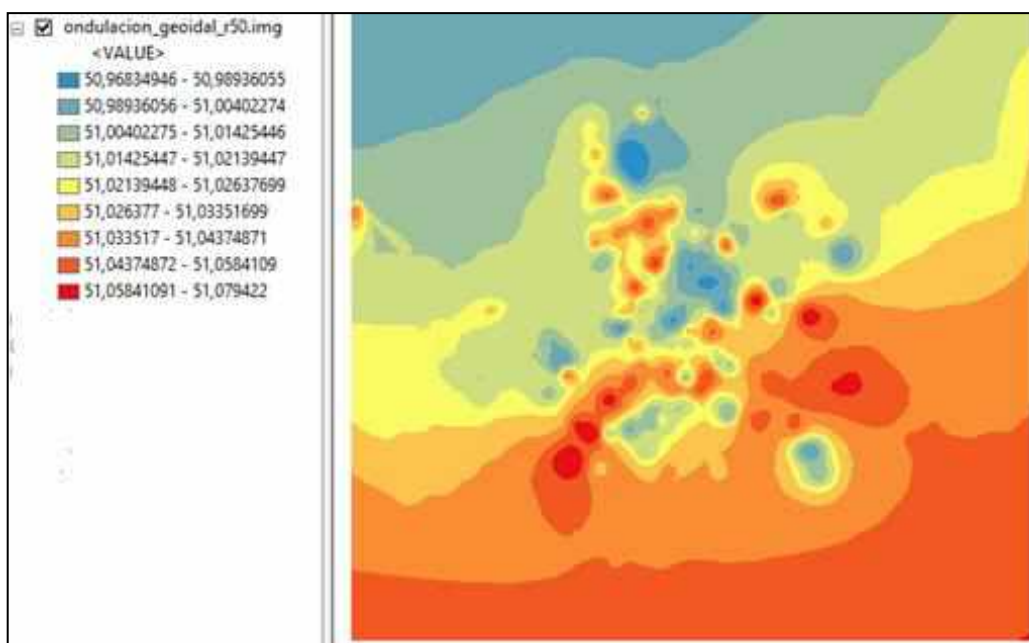


Figura 26: Superficie de estimación de N para el municipio de Madrid, usando ArcGIS (ver el mapa en detalle en el Anexo 4)

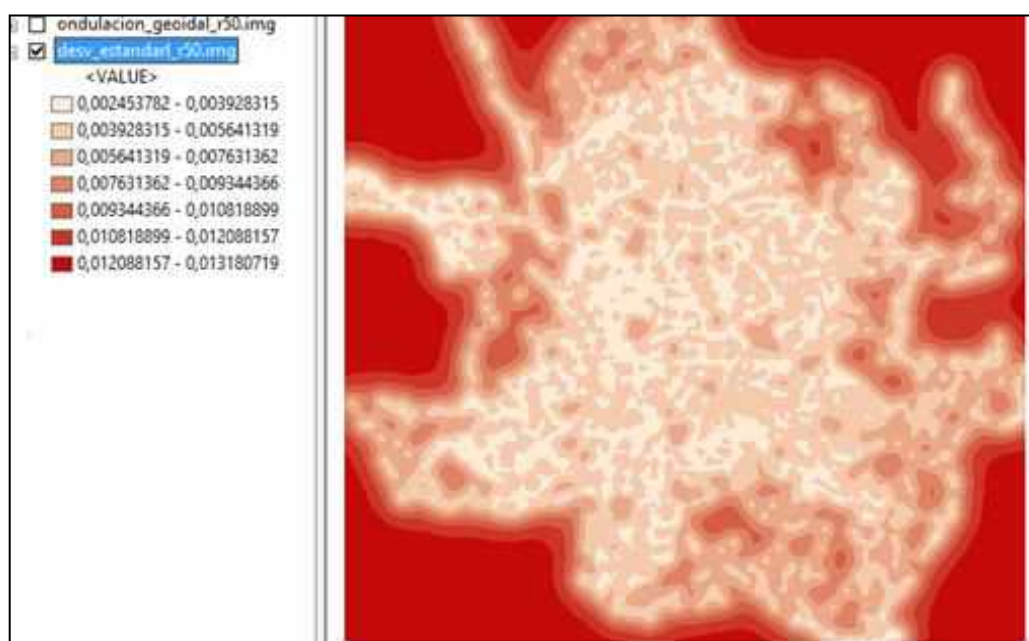


Figura 27: Superficie de estimación de la desviación estándar de N para el municipio de Madrid, usando ArcGIS (ver el mapa en detalle en el Anexo 5)

En el Anexo 4, se observa el mapa del modelo geoestadístico de la Ondulación Geoidal, generada por ArcGIS para el municipio de Madrid. Este mapa indica el valor medio de N para la ciudad, el cual ronda los 51 m. No obstante, N no se mantiene constante sobre la superficie de estudio, ya que va creciendo gradualmente en sentido Norte - Sur a lo largo de todo el límite municipal, con una variación máxima de ± 11 cm. Por otra parte, si se considera la dirección Oeste - Este, se observa que la ondulación geoidal disminuye hacia el centro y se mantiene tenuemente constante en los extremos, con la particularidad que los valores altos de N se concentran al Este de la ciudad.

En razón a lo descrito, se puede afirmar que el comportamiento de la ondulación geoidal para el municipio de Madrid es directamente proporcional a la topografía del terreno. Es decir, a mayor elevación mayor ondulación geoidal y viceversa, como se observa en los distritos de Tetuán, Ciudad Lineal, Moratalaz, San Blas y Carabanchel, los cuales se caracterizan por ser zonas elevadas dentro de la ciudad.

Para finalizar con los resultados obtenidos usando ArcGIS para modelar la variable N, en el Anexo 5 se muestra un mapa con la desviación estándar del error de predicción del modelo. Dicho mapa señala que los errores de estimación de este producto varían entre los 0.25 a 1 centímetros (nivel de confianza del 95%.) en el centro de la municipalidad de Madrid. Sin embargo, la desviación estándar empieza a superar el valor del centímetro en las zonas donde no existió una buena densidad de datos, como fue en los extremos de los distritos de Fuencarral y Barajas.

4.2 MODELO KRIGING USANDO EL SOFTWARE R-STUDIO

Como en el caso anterior, se procedió a dividir los datos originales, previamente depurados, en dos conjuntos aleatorios (independientes al apartado 3.3), con el mismo objetivo de modelar y validar con el software R-Studio la ondulación geoidal del término municipal de Madrid.

R-Studio es un software libre que usa un conjunto de librerías para realizar diferentes cálculos y procedimientos. Las librerías utilizadas e instaladas para poder realizar el presente apartado se enlistan a continuación:

- library (Rcpp)
- library (RGeostats)
- library (Rcmdr)
- library (Rcpp)
- library(ggplot2)
- library(lattice)
- library(caret)

Las funciones derivadas de estas librerías, las cuales fueron utilizadas para cada proceso, se encuentran documentadas en el Anexo 1. Cabe mencionar que, en este apartado solo se expondrá el procedimiento y el análisis de los resultados obtenidos al usar R, ya que la justificación de los parámetros considerados para realizar un modelo Kriging de N para el municipio de Madrid, serán muy similares a los criterios tomados en el apartado anterior (apartado 3.3.)

4.2.1 ANÁLISIS EXPLORATORIO DE DATOS

En la Tabla 4 se muestra, el resumen estadístico generado por R de los valores de la muestra (primer conjunto de datos) de ondulación geoidal para Madrid.

Tabla 4: Resumen estadístico de la variable N, generado por R

| ESTADÍSTICO | VALOR |
|------------------------|--------|
| Media (m) | 51.03 |
| Desv est (m) | 0.016 |
| Coefficiente variación | 0.0003 |
| Valor Min (m) | 50.97 |
| Valor Max (m) | 51.09 |
| 1er Cuartil (m) | 51.01 |
| Mediana (m) | 51.02 |
| 3er Cuartil (m) | 51.04 |
| n | 3152 |

Al igual que en el análisis realizado en ArcGIS, se observa que la distribución de este nuevo conjunto de datos aleatorio es de tendencia central, ya que la media y la mediana toman valores cercanos. La muestra presenta un rango de 0.12 m con un coeficiente inferior al 0.1%, verificando que no existen problemas de variabilidad en los datos. El histograma (figura 28) y el diagrama de caja (figura 29) de los valores de N muestran que la distribución está saturada levemente, hacia la derecha. Además, el diagrama de caja, no muestra observaciones atípicas o valores extremos ya que sus bigotes son relativamente equidistantes a la media.

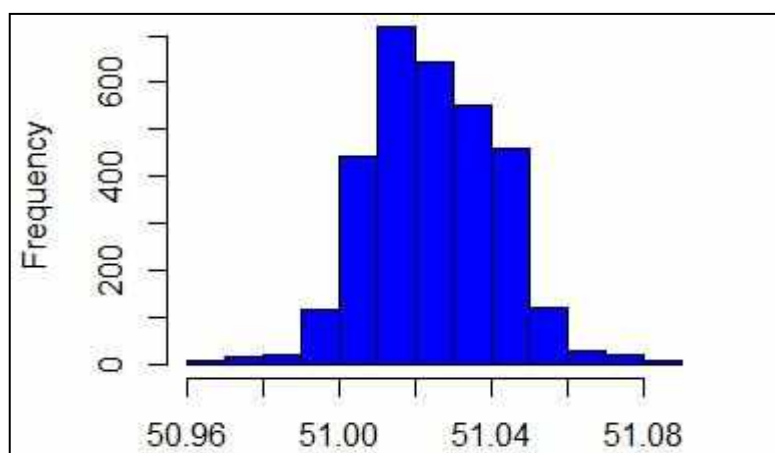


Figura 28: Histograma Ondulación Geoidal – Madrid, en R

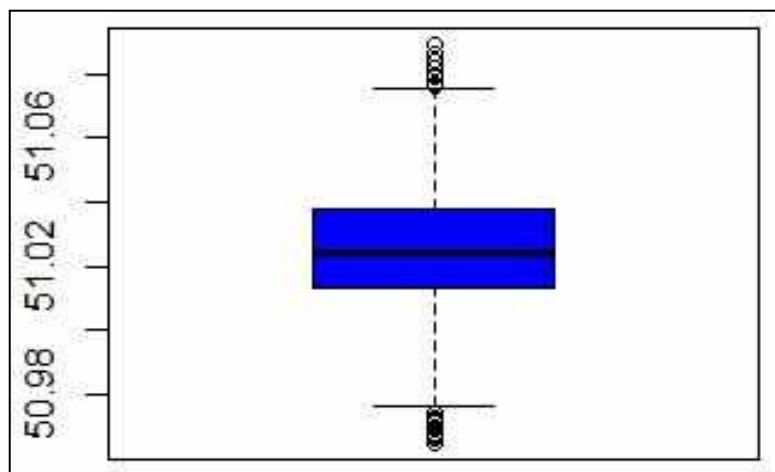


Figura 29: Diagrama de caja de N, en R

El siguiente paso fue analizar la estacionaridad de la variable. En este caso, R genera gráficos de tendencia, separando los componentes de la posición, mientras que, como se vio, ArcGIS lo hace en sus 3 dimensiones (figura 20), siendo esta última más difícil de interpretar por la cantidad de información representada. Al analizar los gráficos de dispersión (figuras 30 y 31 de manera respectiva), se evidencia que el valor de N crece de forma paulatina con respecto al valor de X (prácticamente constante); sin embargo, aproximadamente en la mitad del dominio de X, se observa como N decrece ligeramente para luego seguir aumentando. Esta zona de decrecimiento corresponde al centro de Madrid. Por su parte, la dependencia de N con respecto a su componente en Y se mantiene inversamente proporcional en toda la ciudad.

Con esta examinación se puede concluir que el proceso aleatorio que describe la ondulación geoidal es no estacionario, debido a que la media y la varianza no son constantes a lo largo de toda la zona de estudio. Sin embargo, podría considerarse estacionaria solo en su componente X, mas no su posición vectorial XY (espacial).

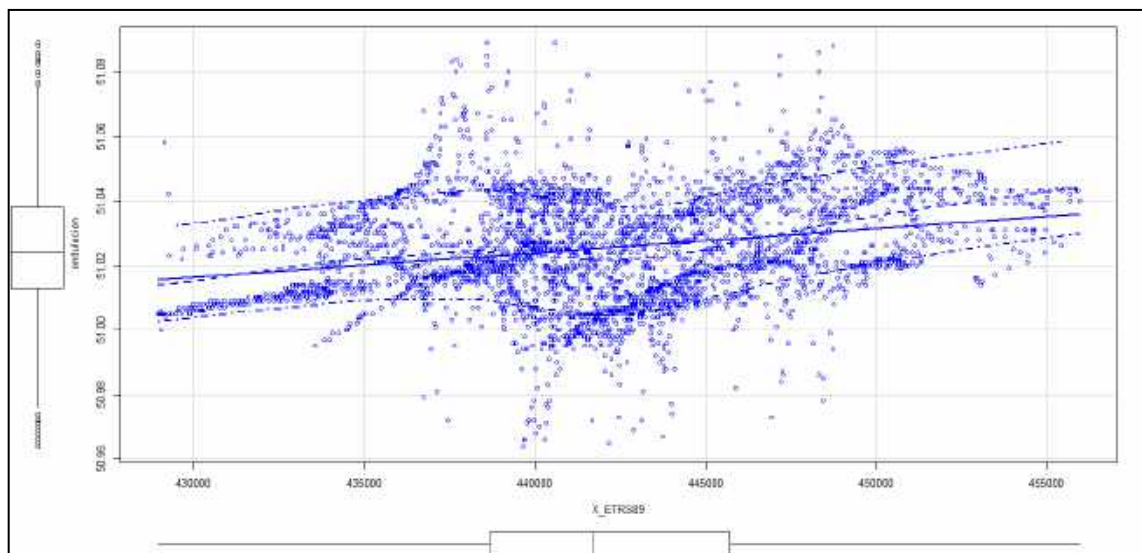


Figura 30: Estacionariedad de la variable en X

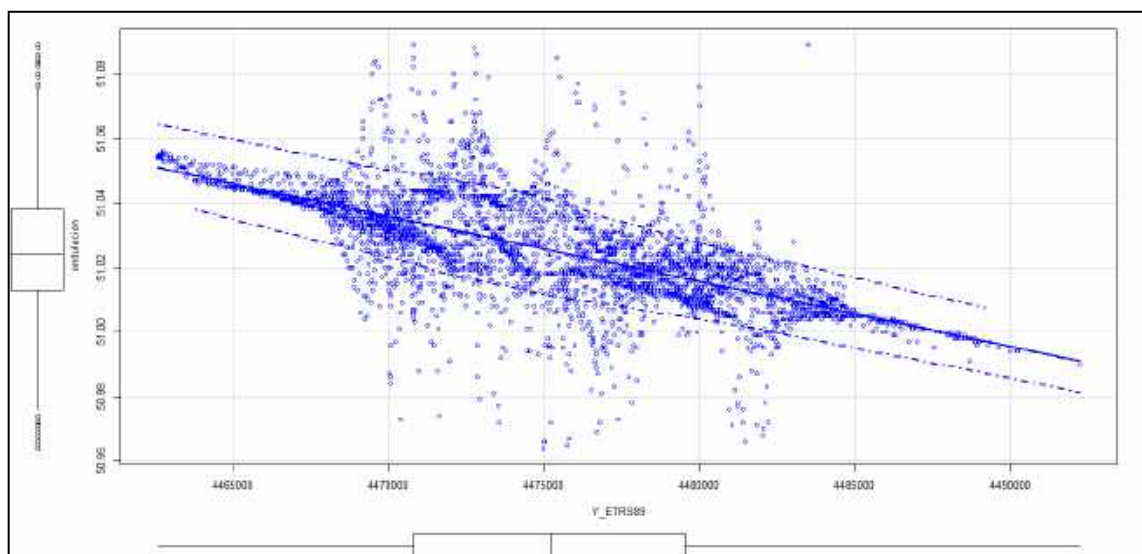


Figura 31: Estacionariedad de la variable en Y

Finalmente, en la Figura 32 se muestra la distribución espacial en función a la magnitud de la variable. Como se determinó en los gráficos anteriores, la magnitud disminuye de manera considerable conforme los vértices de RTM se distribuyen hacia el norte; no obstante, la variable en sentido Este – Oeste crece a una razón de cambio muy bajo.

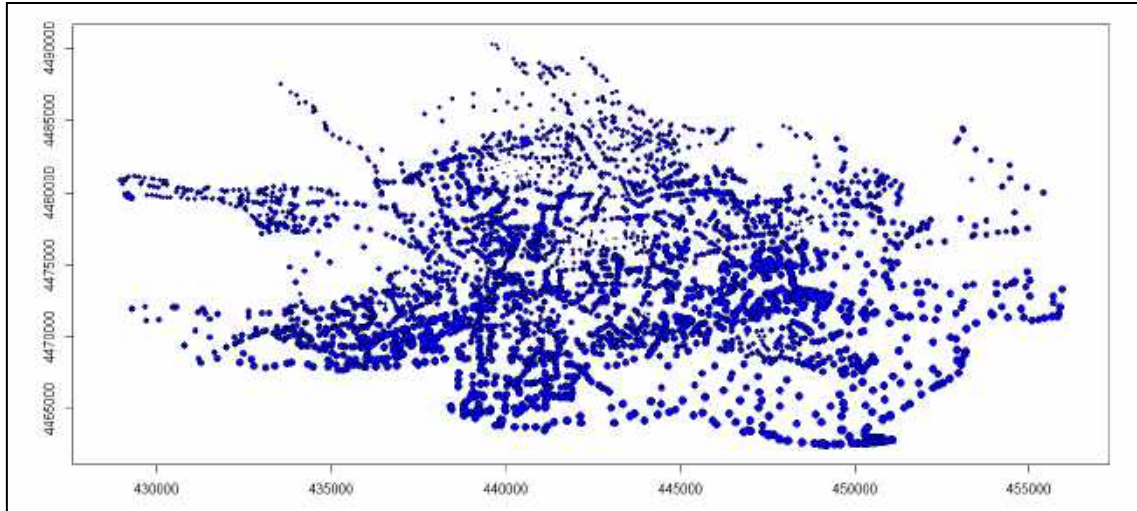


Figura 32: Distribución espacial de datos

4.2.2 VARIOGRAMA EXPERIMENTAL

Tal como se explicó en el apartado anterior es importante definir si un variograma estudiado depende de la dirección de su análisis, por ello se generaron 4 variogramas direccionales con las siguientes características:

- variograma 0° azul
- variograma 45° rojo
- variograma 90° verde
- variograma 135° negro

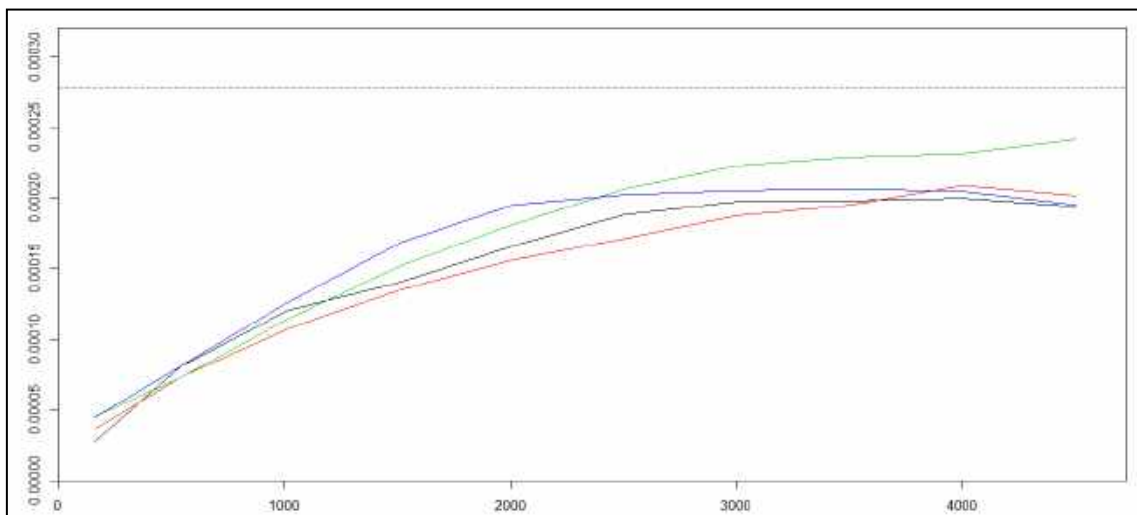


Figura 33: Variogramas direccionales generado en R

En la Figura 33 se puede observar que R representa de mejor manera los variogramas direccionales, ya que utiliza líneas de tendencia y no una nube de dispersión como lo hace ArcGIS. Esto facilita la interpretación, y por ello se decidió analizar el variograma

experimental con Lag más distantes (500 m), con el objetivo de afirmar cuál es el valor del rango donde la función empieza acotarse, es decir hasta donde existe dependencia por correlación.

Como en el subapartado 4.1.2, al comparar los 4 variogramas direccionales se determinó que estos no presentan una variación significativa en función a la dirección del análisis, por lo tanto, se trata al variograma experimental como uno omnidireccional (isotrópico). Para la construcción del variograma experimental omnidireccional se usó la librería RGeostats de R, con la función *vario.calc* (ver Anexo 1). Para esto se utilizaron 3152 vértices de la RTM de Madrid, con un rango en el variograma aproximado de 3000 m (valor donde la función se acota), el cual fue definido en función de 10 Lags con un valor de 300 m.

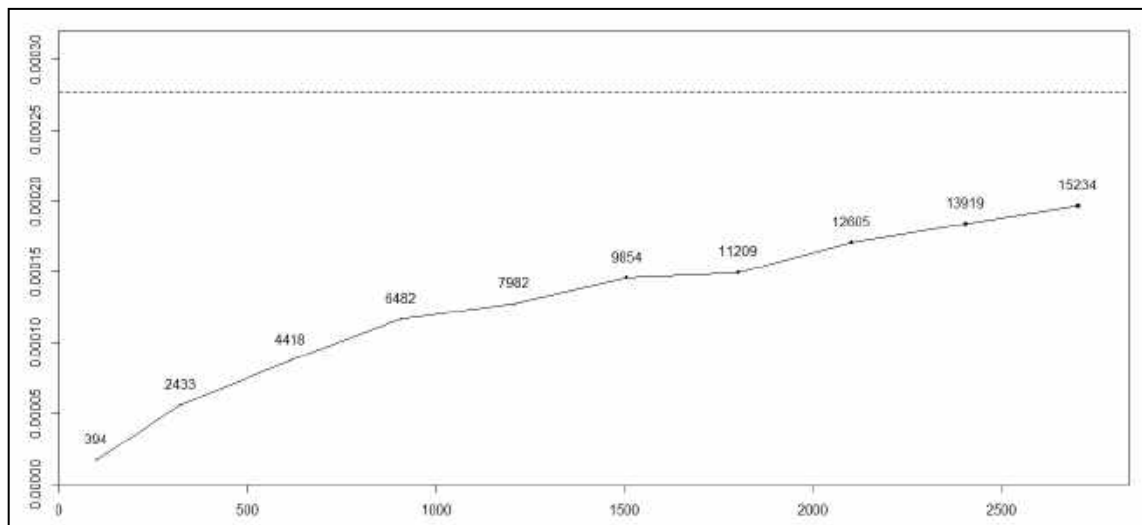


Figura 34: Variograma experimental omnidireccional generado en R

Con lo antes mencionado, al examinar la Figura 34 se extrajeron las siguientes características del variograma experimental:

- Presenta efecto pepita (Nugget Effect) aproximado de 0.00002 m^2
- El valor de alcance (Rango) aproximado es de 2750 m.
- El valor de la meseta (Sill) aproximada es de 0.00015 m^2 .
- Por la forma del variograma experimental, los modelos teóricos que podrían ajustarse a esta estructura son: exponencial, gaussiano, esférico, stable y K-bessel (ver Anexo 2).

Como se puede observar, los parámetros de la meseta, alcance y efecto pepita de este variograma, así como los posibles modelos de ajuste del mismo, difieren del apartado 4.1.3. Esto se debe a que las muestras, el modelo y el análisis de estructura en R es completamente independiente al realizado en ArcGIS, a pesar de utilizar los mismos datos de entrada.

4.2.3 NUBE VARIOGRÁFICA

La denominada nube variográfica expresa de manera gráfica la semivarianza de dos puntos de la variable en estudio, en función de la distancia existente entre ellos (Alfaro, 2007). Según Leiva (2014), esta herramienta permite identificar la influencia de valores atípicos en el cálculo del variograma experimental. Esta nube fue generada con la función “cloud.cal” del paquete RGeostat y sus resultados se presentan en la figura 35. De su análisis se concluye que no existen valores atípicos, debido a su bajísima variabilidad

(coeficiente de variación menor a 0.1%), observando, además, que la semivarianza se mantiene constante y bien cohesionada en distancias de 5 a 10 km (colores rojizos).

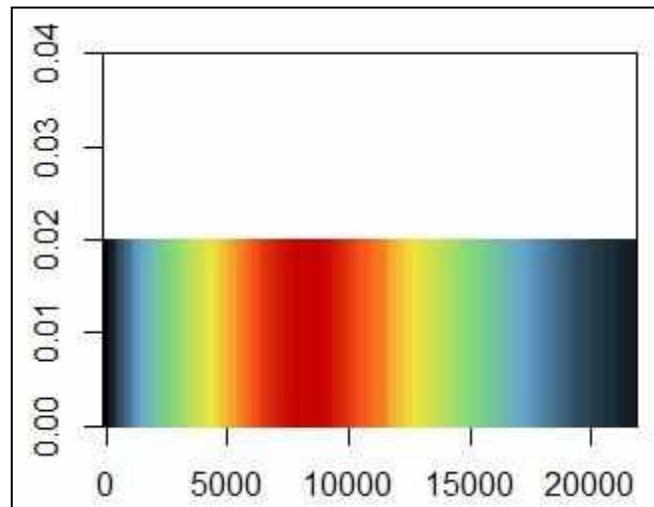


Figura 35: Nube variográfica generada en R

4.2.4 VARIOGRAMA TEÓRICO

En el presente apartado se evaluaron 5 posibles funciones de ajuste para el variograma experimental. Para llevarlo a cabo se utilizó la función *auto.model* de la librería RGeostats, que permite ingresar los posibles modelos teóricos (estructuras básicas) simultáneamente, puesto que R intenta minimizar estos en una sola función de costo. Leiva (2014) define esta función como “la diferencia en todos los puntos entre el variograma experimental y los modelos de ajuste, minimizándose a través del algoritmo de Gauss-Newton”. En otras palabras, el ajuste resultante será una función anidada de estructuras básicas que reducen al mínimo los valores residuales de la estimación.

Las funciones evaluadas para el ajuste fueron: exponencial, gaussiano, esférico, stable y K-bessel (Anexo 1) y el efecto pepita. Como resultado se obtuvo una función anidada, en la que se descartaron los modelos esféricos, k-Bessel y la estructura efecto pepita. Este resultado se observa mejor en la Tabla 5 y en la Figura 36.

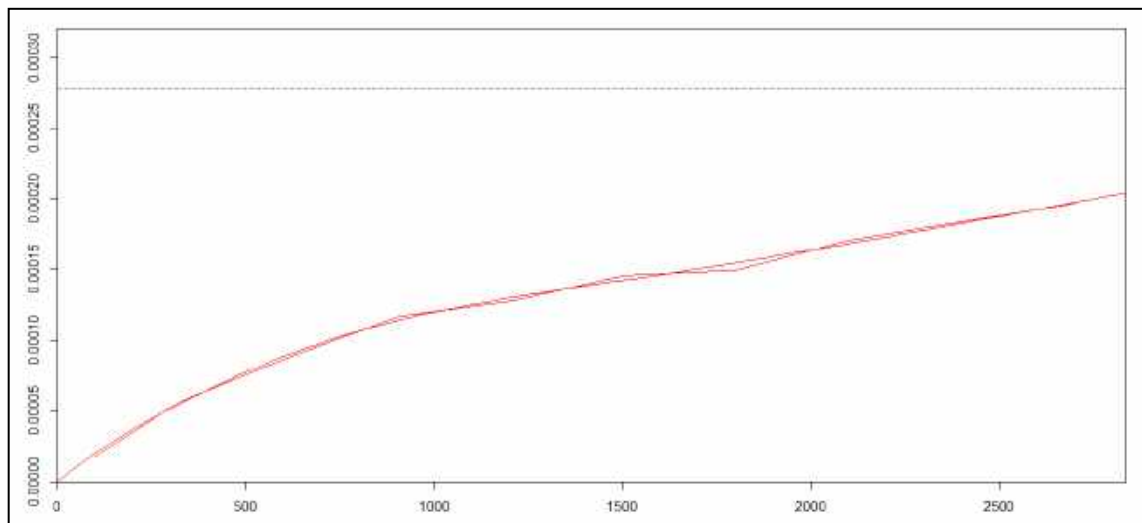


Figura 36: Ajuste de variograma teórico de N en R

Tabla 5: Parámetros de Estructuras Básicas - R

| Modelos Básicos | Parámetros | |
|-----------------|------------|------------------------|
| | Rango (m) | Sill (m ²) |
| Exponencial | 780 | 5.92e-05 |
| Gaussiano | 1140 | 5.31e-05 |
| Stable | 6901 | 0.000231 |

Debido a la facilidad de modelar el variograma teórico que posee la aplicación R, se realizaron algunas pruebas con la función *auto.model* para comprobar resultados. Por ejemplo, en la Figura 37, se ajustó el variograma experimental, usando las mismas funciones anidadas que se obtuvieron de ArcGIS (Gaussiano + circular o esférico), mientras que en la Figura 38, se ajustó solo considerando el modelo esférico.

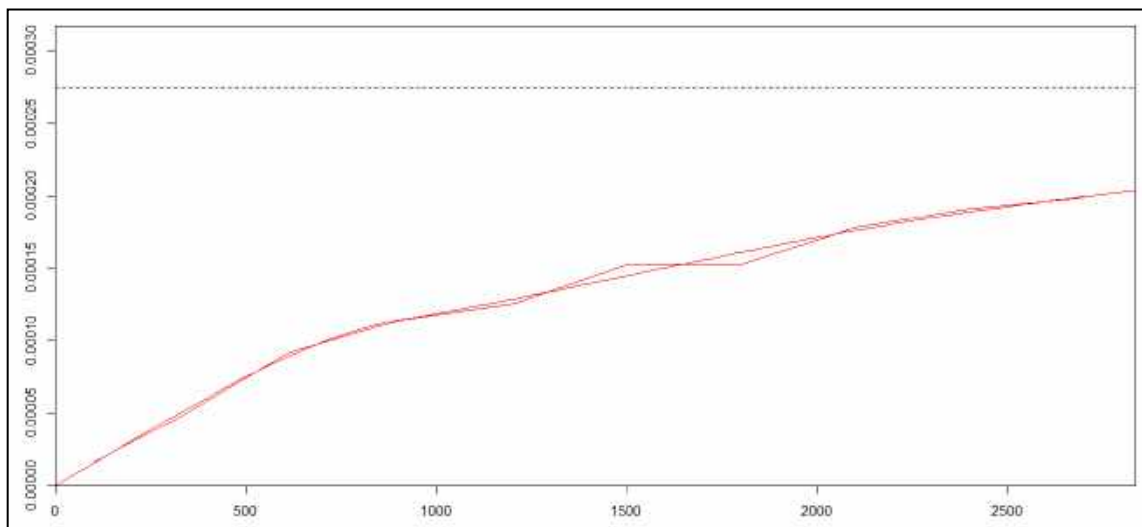


Figura 37: Ajuste de variograma teórico con el mismo modelo anidado de ArcGIS

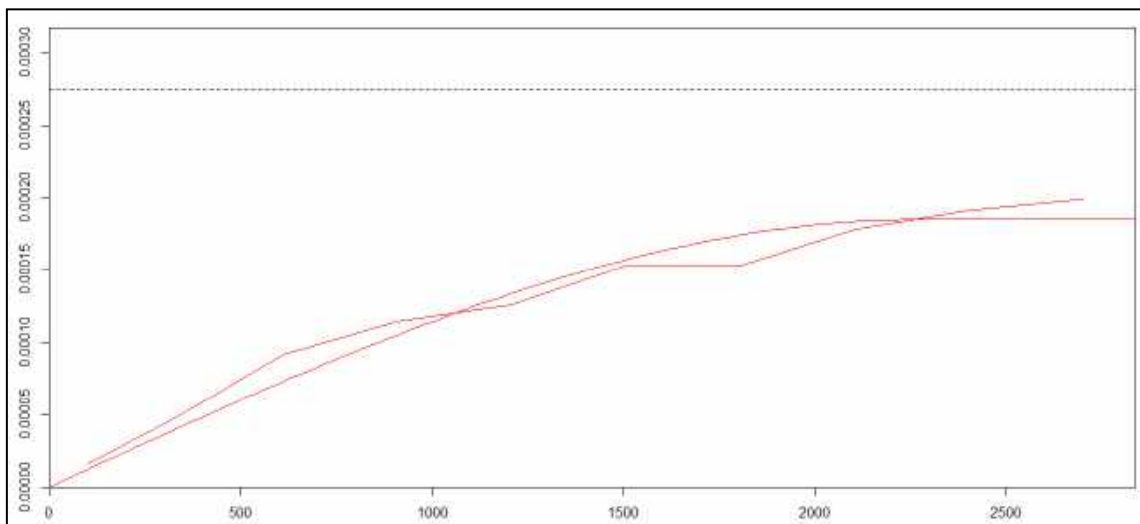


Figura 38: Ajuste de variograma teórico solo con modelo esférico.

Evidentemente, el modelo que mejor se ajusta al variograma experimental es el primero (figura 36), pero cabe destacar que las diferencias con el modelo anidado usado en ArcGIS son mínimas.

4.2.5 VALIDACIÓN CRUZADA INTERNA

Al igual que las funcionalidades de ArcGIS, R también necesita de una matriz de búsqueda para realizar la validación cruzada interna. Esta matriz no se genera de manera automática como lo define ArcGIS, por lo tanto, para R es necesario definir de manera precisa cada parámetro de búsqueda, siendo estos los siguientes:

- Tipo de búsqueda: Vecindad móvil (2), solo seleccionará los datos activos ubicados en las proximidades del objetivo (R Documentation)
- Mínimo número de muestras en la vecindad: 8, se ha demostrado que, para variables de dos dimensiones, 8 muestras dan buenos resultados (Alfaro, 2007).
- Máximo número de muestras en la vecindad: 16, Oliver y Webster (2010) recomiendan que la vecindad debe estar en un máximo de 16 a 20 muestras.
- Definición de sectores: Sí (para desagregación de las muestras)
- Número de sectores: 8 (Octantes), recomendado por Leiva (2014)
- Máximo número de muestras por sector: 2
- Máxima distancia de búsqueda: 3000 m en función del variograma experimental y teórico, ver figura 36.

Con los parámetros antes mencionados se creó la matriz de búsqueda utilizando la función *neigh.input* de la librería RGeostat (ver Anexo 3). De igual manera, empleando la misma librería se usó la función *xvalid* para ejecutar la validación cruzada interna (ver Anexos 1 y 3), obteniendo como resultado un histograma de los errores de dicha validación (figura 39). El resultado muestra que la mayor concentración de datos está dentro de un intervalo inferior a ± 0.01 m, ratificando la buena correspondencia que existe entre los datos, el ajuste del modelo y la matriz de búsqueda. Sin embargo, existen algunos residuales que se encuentran por encima del intervalo ± 0.025 m, lo que podría ser sinónimo de heterocedasticidad, como se determinó también en el subapartado 4.1.4 usando el software ArcGIS.

Finalmente, con la nube de residuales generada por R (3149 muestras) se calculó el RMSE, obteniendo un valor de 6.2 mm.

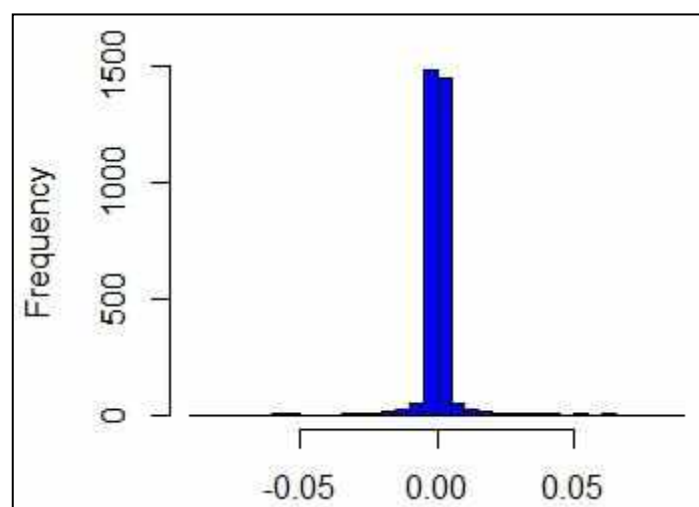


Figura 39: Error de validación cruzada interna por R

4.2.6 SUPERFICIE DE PREDICCIÓN

La modelización cartográfica de la superficie de predicción espacial de la ondulación geoidal para el término municipal de Madrid, utilizando el software libre R-Studio, se realizó a través de un modelo estocástico de estimación Kriging ordinario puntual, que se exportó, igualmente, sobre una malla regular de puntos con intervalo de 50 metros en un fichero CSV y ASC. La selección del intervalo de puntos fue preferida bajo el mismo criterio expuesto en el subapartado 4.1.5

La estimación espacial se generó con la función kriging de la librería RGeostats de R, con los parámetros definidos al obtener el variograma teórico (ver subapartado 4.1.4), para una vecindad móvil. Como resultado final se obtuvo una nube de puntos con los valores de predicción de N para el municipio de Madrid, además de su respectiva desviación estándar. En la Figura 40 el software R intenta representar espacialmente la superficie de estimación, sin embargo, la representación espacial y gráfica de una variable no es una fortaleza de este software. Por lo tanto, la nube de puntos resultante en R se convirtió a formato raster para, posteriormente, ser mapeado en ArcGIS, con el objetivo de representar y visualizar mejor la variable (Anexo 6 y Anexo 7).

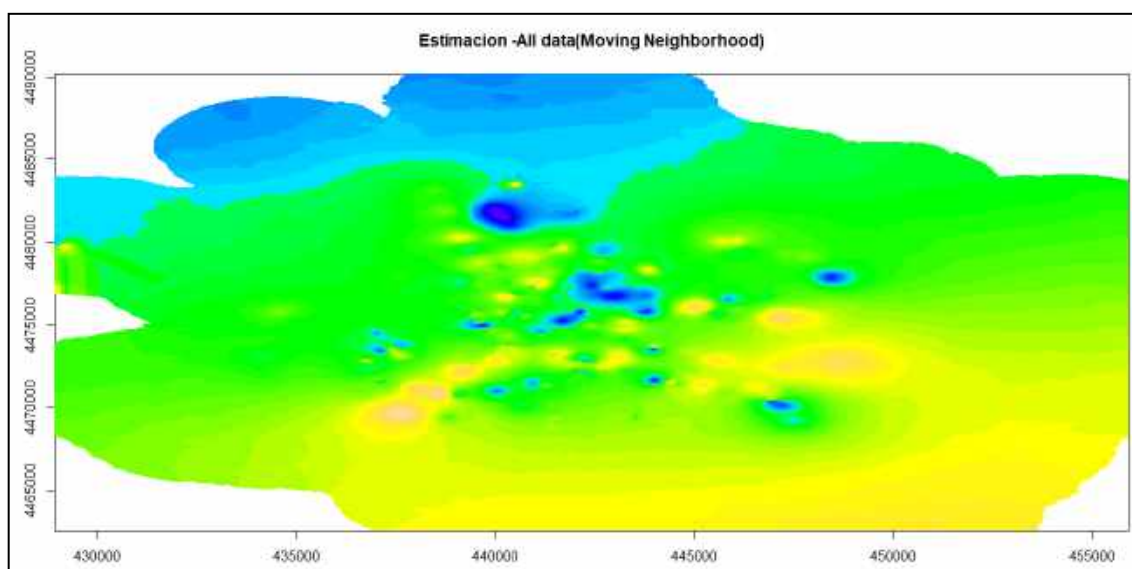


Figura 40: Superficie de estimación de N para la ciudad de Madrid, usando R (ver mapa en el Anexo 6)

En el Anexo 6, se observa el mapa del modelo geoestadístico de la Ondulación Geoidal, generada por R-Studio para el municipio de Madrid. Este mapa, presenta un comportamiento muy similar al explicado al final del subapartado 4.1.5. A pesar de ello, el modelo generado por R posee un rango más amplio que el descrito por su homólogo en ArcGIS. Motivo por el cual, su mapa (Anexo 6) presenta mayor variación de tonalidades azules (corresponde con valores bajos) al norte de la zona de estudio, sucediendo lo mismo con las tonalidades anaranjadas (valores altos) al límite sureste de la ciudad. Por lo tanto, ya que las diferencia entre mapas son netamente estéticas por la diferencia entre rangos, se determinó que ambos modelos describen el comportamiento de la ondulación geoidal en función de la topografía para el municipio de Madrid.

En el Anexo 7 se puede observar el mapa con la desviación estándar del error de predicción del modelo generado por R-Studio, el cual como en el caso anterior, posee un

comportamiento similar a su homólogo en ArcGIS. La diferencia radica en que este modelo varía entre los 0.05 a 1.55 cm (nivel de confianza al 95%.) en el centro de Madrid y con una desviación que puede superar los 2 cm en los extremos occidentales del distrito de Fuencarral.

4.3 VALIDACIÓN CRUZADA EXTERNA: COMPARACIÓN DE LOS MODELOS KRIGING ARCGIS Y R-STUDIO CON EL MODELO EGM08 – REDNAP (IGN)

Como se mencionó en el capítulo anterior, los programas utilizados para el presente TFM generaron una validación cruzada interna, que utilizaban los mismos datos empleados para modelar y validar la superficie de estimación. Sin embargo, Giraldo y Cañadas (2007, 2013) recomiendan que dicho proceso sea realizado con datos diferentes. Por lo tanto, en el presente apartado se detalla el manejo de datos independientes aleatorios (previamente separados), correspondientes a un 20% de la muestra total, que no fueron incluidos en la creación de las superficies de estimación anteriormente modeladas.

En este trabajo este proceso se denomina como validación cruzada externa. La diferencia entre ambos procesos (validación cruzada interna y validación cruzada externa) se fundamenta en que la validación externa no requiere una matriz de búsqueda, pues es suficiente con la extracción y comparación entre los valores estimados con sus correspondientes valores observados (residual). Para la extracción de los valores estimados de N se utilizó la herramienta *Extract value to point* de ArcGIS, la cual permite obtener el valor de un punto ubicado sobre una superficie o capa ráster (ESRI, 2019). Posteriormente con un campo ID (identificador) se ordenaron los datos de forma agrupada para calcular su respectivo residual.

Por otro lado, para obtener una apreciación adicional en la calidad y validación de los resultados del TFM se descargó del Geoportal del IGN el modelo de Ondulación Geoidal EGM08 – REDNAP, pues este es de carácter nacional y oficial para España, con el objetivo de compararlo con los mismos puntos validados de ambos modelos Kriging desarrollados, obteniendo los siguientes resultados, los cuales fueron valorados en función a su precisión utilizando el RMSE (Tabla 6):

Tabla 6: Validación cruzada externa de modelos Kriging VS modelo IGN

| VALIDACIÓN CRUZADA EXTERNA | | | | |
|--|----------------|----------------|------------------|----------------|
| MODELOS DE ESTIMACIÓN | Kriging ArcGIS | EGM08 – REDNAP | Kriging R-Studio | EGM08 – REDNAP |
| Muestras | 786 | 786(1) | 784 | 784(2) |
| N Estimado (m)* | 51,025 | 51,127 | 51,026 | 51,127 |
| 1 Desv. Estándar (cm)* (68% de confianza) | 0,450 | - | 0,604 | - |
| 2 Desv. Estándar (cm)* (96% de confianza) | 0,900 | - | 1,208 | - |
| RMSE (cm) | 0,449 | 10,635 | 0,463 | 10,532 |

*Valores medios respectivos

(1) Muestras usadas para validar en ArcGIS

(2) Muestras usadas para validar en R-Studio

4.4 PRECISIÓN DE MODELOS

Como se ha visto en los capítulos precedentes, se han analizado diferentes formas de validar y comprobar la calidad de los modelos de estimación de la Ondulación Geoidal para la práctica totalidad del municipio de Madrid, cuyo objetivo ha sido determinar qué superficie de estimación generada en el presente TFM describe de mejor manera el comportamiento de la variable de estudio.

En la Tabla 7 se presenta un resumen de los resultados obtenidos, donde se ilustra como cada modelo posee errores dentro de los límites de precisión esperada (desviación estándar). Esto se reitera al analizar el valor RMSE de cada modelo: ninguno superó el valor de su respectiva desviación estándar. Por otra parte, en la misma tabla se puede evidenciar que la diferencia entre ambos modelos generados, es mínima (inferiores a 1 milímetro), ya sea comparando cualquiera de sus validaciones cruzadas.

Sin embargo, si se coteja el RMSE de la validación interna del modelo ArcGIS, con su respectivo par del modelo R-Studio, se observa una diferencia de 0.73 mm, mientras que para la validación externa su valor es de 0.14 mm. Leiva (2014) afirma que “si el modelo de semivarianza elegido describe bien la estructura de autocorrelación espacial, la diferencia entre el valor observado y el valor predicho debe ser mínimo”. Por ende, bajo esta premisa, se tomó al modelo Kriging generado por ArcGIS como la superficie que mejor describe la ondulación Geoidal en la ciudad de Madrid. Por otra parte, debido a que las precisiones entre ambos modelos locales son mínimas, se debe enfatizar que el software comercial optimiza el tiempo de trabajo, al poseer procesos automatizados, pero el uso de sus licencias incrementa los costos de producción. Mientras que, el software libre abarata costos y permite libertad en la manipulación de procesos, pero su tiempo de producción es más lento.

Tabla 7: Resumen de precisiones alcanzadas por cada modelo analizado

| RESUMEN PRECISIONES | | | | | |
|------------------------------|-----------------------|----------------|-------------------------|------------|-----------------------|
| MODELOS DE ESTIMACIÓN | Kriging ArcGIS | | Kriging R-Studio | | EGM08 – REDNAP |
| Método evaluación | VCI (1) | VCE (2) | VCI | VCE | VCE |
| RMSE (cm) | 0,553 | 0,449 | 0,626 | 0,463 | 10,583* |
| 1 Desv. Est (cm)* | 0,818 | 0,450 | 0,896 | 0,604 | - |

*Valores medios respectivos de cada modelo

(1) Validación cruzada interna

(2) Validación cruzada externa

Finalmente, como se enunció en el capítulo anterior, se utilizó muestras de la validación cruzada externa de ambos modelos para calcular el error medio cuadrático existente en el modelo EGM08 – REDNAP del IGN. Dicho error es poco más de 10 cm, mientras que el error máximo entre ambos modelos kriging no superaron el medio centímetro, es decir aproximadamente 25 veces más preciso (valores de VCE). Por lo cual, se puede afirmar que ambas superficies de estimación de la variable N son más precisas para la ciudad de Madrid, que su homólogo de carácter nacional.

No obstante, hay que mencionar también que los datos usados en este trabajo (red RTM) poseen una densificación de un vértice cada 250 metros, mientras que los datos usados para el modelo EGM08 – REDNAP tienen una densidad de un vértice cada 5 km (Instituto Geográfico Nacional, 2009). En consecuencia, al realizar una proporción de la densidad de los datos de cada red, se puede ver que la RTM es 20 veces más densa que la REDNAP, debido a que la primera cubre sólo el entorno de la ciudad de Madrid mientras que la segunda a toda la península Ibérica.

Por otra parte, el Instituto Geográfico Nacional generó su modelo EGM08 – REDNAP con técnicas de interpolación determinísticas basadas en algoritmos que minimizan los residuales al máximo sobre puntos observados (teoría de la elasticidad) (Instituto Geográfico Nacional, 2009). El problema de utilizar algoritmos como el descrito, es que estos fuerzan a la superficie de estimación a ajustarse a aquellos puntos que poseen datos, mientras que las zonas no densificadas pueden sufrir de una exagerada sobrestimación (Desassis y Renard, 2013), disminuyendo así la precisión y estabilidad del modelo. Mientras que la interpolación por métodos estocásticos como kriging generan un ajuste medio para toda la superficie delimitada por los datos base (puntos observados), en función a la semivariancia y a la dependencia espacial, creando una distribución lo más homogénea posible de sus errores (Giraldo, 2007). Por lo tanto, es evidente que parte de la alta precisión obtenida en el presente documento tuvo como eje la calidad y distribución de los datos base, así como el método utilizado para la interpolación, ya que para redes tan densas que están espacialmente correlacionadas siempre es mejor utilizar métodos estocásticos a determinísticos (Emery, 2013).

4.5 DISTRIBUCIÓN DE ERRORES

Como se mencionó a lo largo de los apartados de la modelización Kriging, tanto usando ArcGIS como R-Studio, se consideró la posibilidad de la existencia de heteroscedasticidad en los dos modelos generados. Esta posibilidad fue ratificada por la validación cruzada interna y externa, mismas que analizaron los valores residuales y pudo determinar que estos presentan una tendencia no homogénea a lo largo de la superficie de estimación. Por lo tanto, es imperativo establecer si este fenómeno es significativo sobre las diferentes superficies de estimación Kriging.

Para valorar dicha significancia se aplicó el Índice de Moran a los residuales de cada modelo, al considerar que este mide la autocorrelación espacial en función a la posición espacial y los valores de su vecindad (ESRI, 2019). En este sentido, dado un conjunto y atributo asociado (residuales para este caso), el índice de Moran evalúa si el patrón expresado está agrupado, disperso o es aleatorio (ESRI, 2019), permitiendo así identificar las zonas y valores que provocan heteroscedasticidad en una superficie. Es preciso señalar, que el índice de Moran calcula una puntuación Z (valores normalizados) y un valor P (área bajo la curva) para evaluar la significancia estadística del patrón analizado.

4.5.1 TENDENCIA DE RESIDUALES DE LA SUPERFICIE KRIGING MODELADA EN ARCGIS

En este subapartado se aplicó el índice de Moran para los residuales obtenidos de la validación cruzada interna y externa del modelo generado en ArcGIS. Es importante mencionar que este análisis es muy susceptible al tipo de vecindad asociada, por ello, se recomienda usar una distancia euclidiana agrupada por una banda automática, que es sugerida por el software (radio de búsqueda).

En la Figura 41 se evidencia justo el cálculo del índice de Moran para la validación cruzada interna, con una banda euclidiana de 1.1 km de influencia. Con este se obtiene un valor normalizado Z de -5.58, que determina la existencia de heteroscedasticidad en zonas dispersas sobre la superficie de estimación, con una significancia estadística menor al 1%.

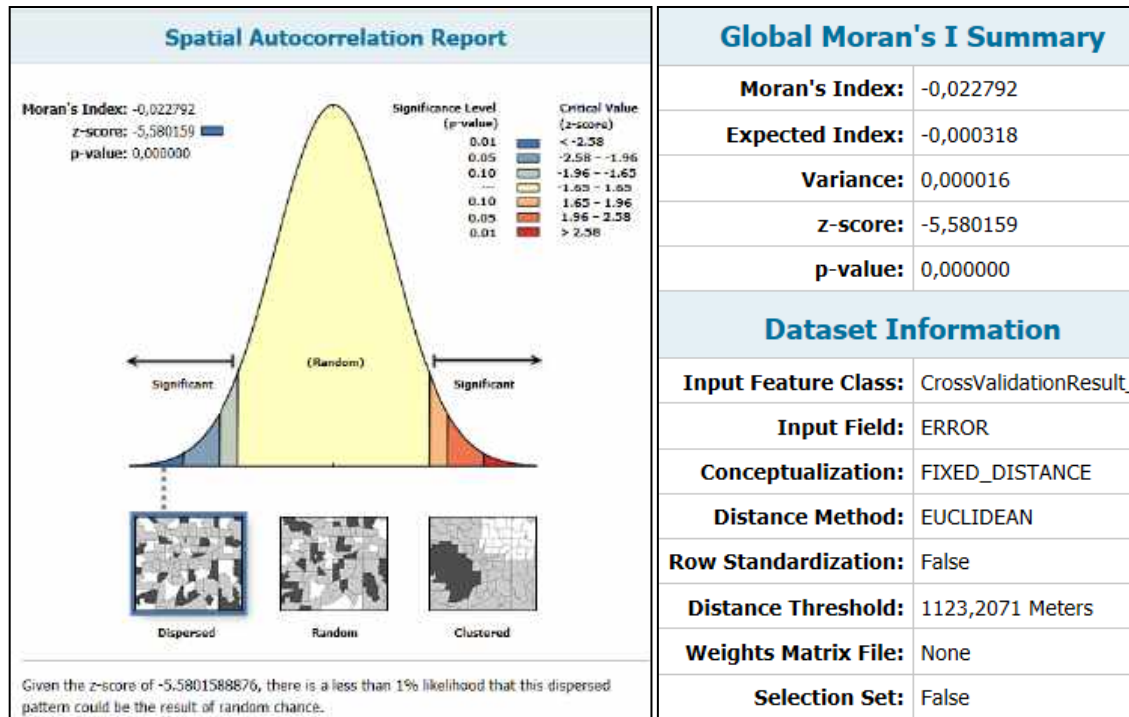


Figura 41: Índice de Moran para validación cruzada interna (modelo ArcGIS)

Sin embargo, en la Figura 42 se observa una discrepancia de resultados, ya que índice de Moran aplicado sobre la validación cruzada externa, con una banda euclidiana de 2.4 km, arrojó un valor normalizado Z de -1.22, lo que indica una distribución homogénea (aleatoria) de los residuales sobre la superficie de estimación a lo largo de la zona de estudio. En otras palabras, según la validación externa el modelo kriging generado por ArcGIS, no presenta heterocedasticidad.

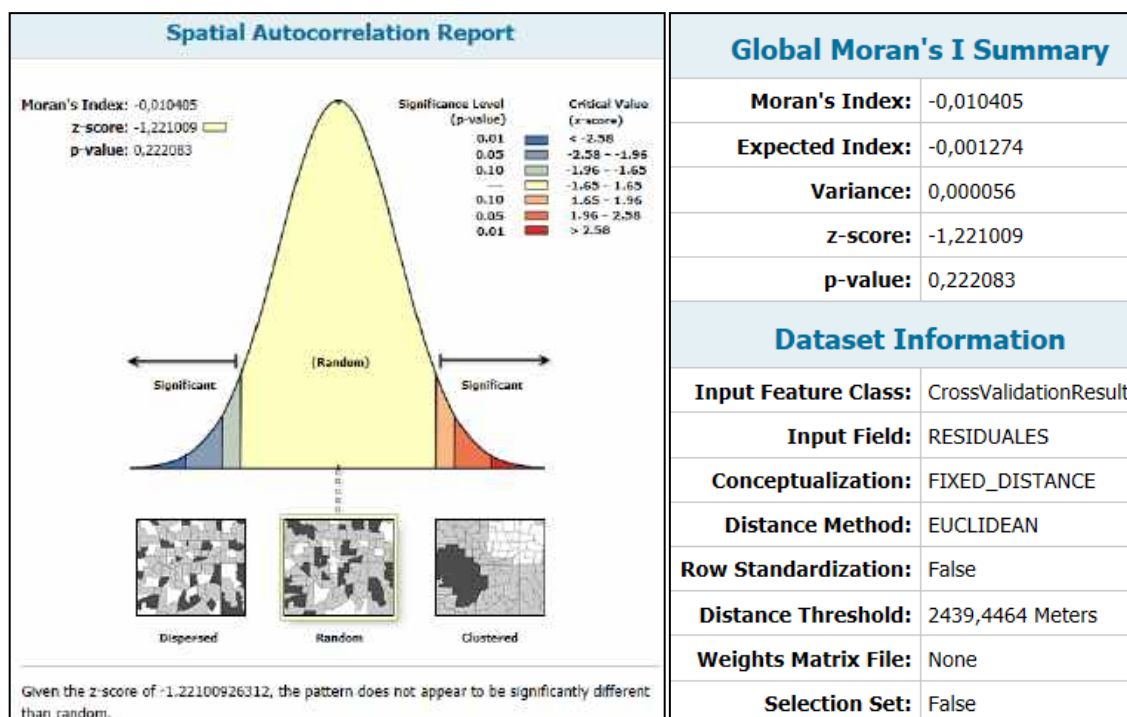


Figura 42: Índice de Moran para validación cruzada externa (modelo ArcGIS)

4.5.2 TENDENCIA DE RESIDUALES DE LA SUPERFICIE KRIGING MODELADA EN R-STUDIO

Como en el caso anterior, se aplicó el índice de Moran a los residuales de ambas validaciones cruzadas del modelo kriging obtenido en R-Studio, siendo sus resultados los que se explican a continuación:

La Figura 43 muestra el índice de Moran de la validación cruzada interna, con una banda euclidiana de 1.1 km (generada por defecto), siendo esta similar al apartado anterior. Este suceso se da por la dispersión de los datos a lo largo de la zona de estudio, ya que son similares, pero no iguales, al ser ambas muestras aleatorias e independientes entre sí. Por lo tanto, al analizar el valor normalizado Z, en este caso de -6.15 (cercano a su homólogo en ArcGIS), se comprobó que la validación cruzada interna determina la existencia de heteroscedasticidad para ambos modelos en zonas dispersas con una significancia estadística menor al 1%.

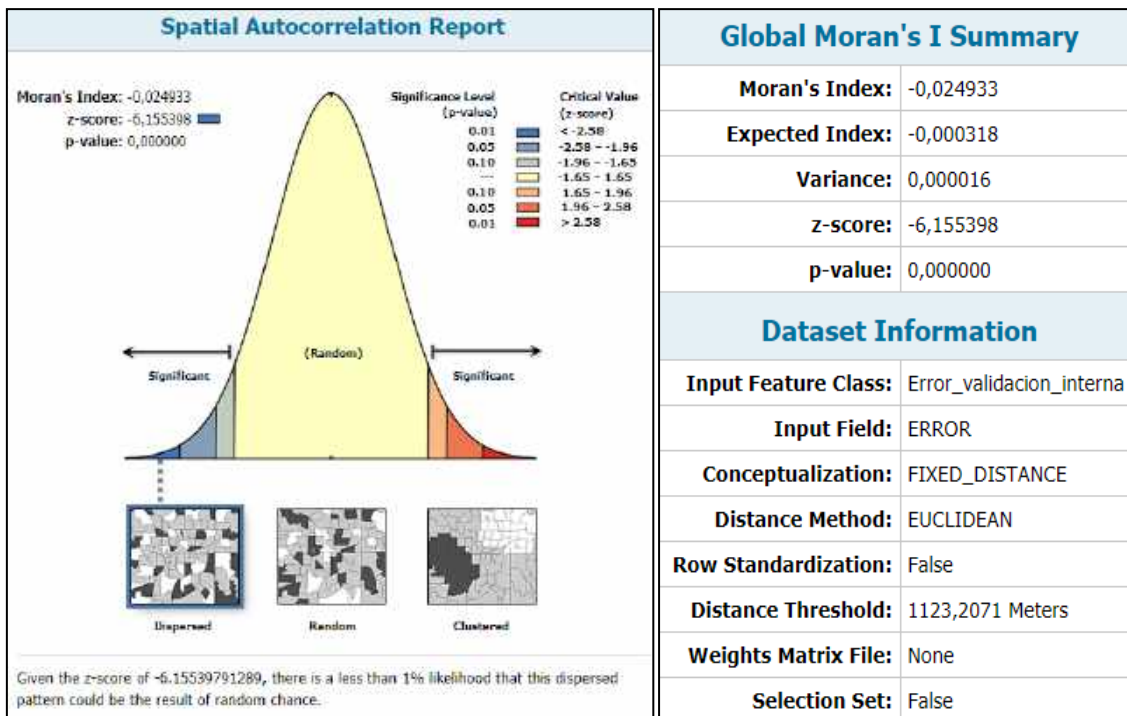


Figura 43: Índice de Moran para validación cruzada interna (modelo R)

Sin embargo, como en el apartado anterior, la Figura 44 permite observar una vez más la discrepancia de resultados entre la validación cruzada interna y externa. Así, el índice de Moran aplicado sobre la validación cruzada externa, con una banda euclidiana de 2.5 km (una vez más similar a su homologo), obtuvo un valor normalizado Z de 1.79, lo que evidencia una distribución conglomerada de residuales con una significancia estadística del 7%.

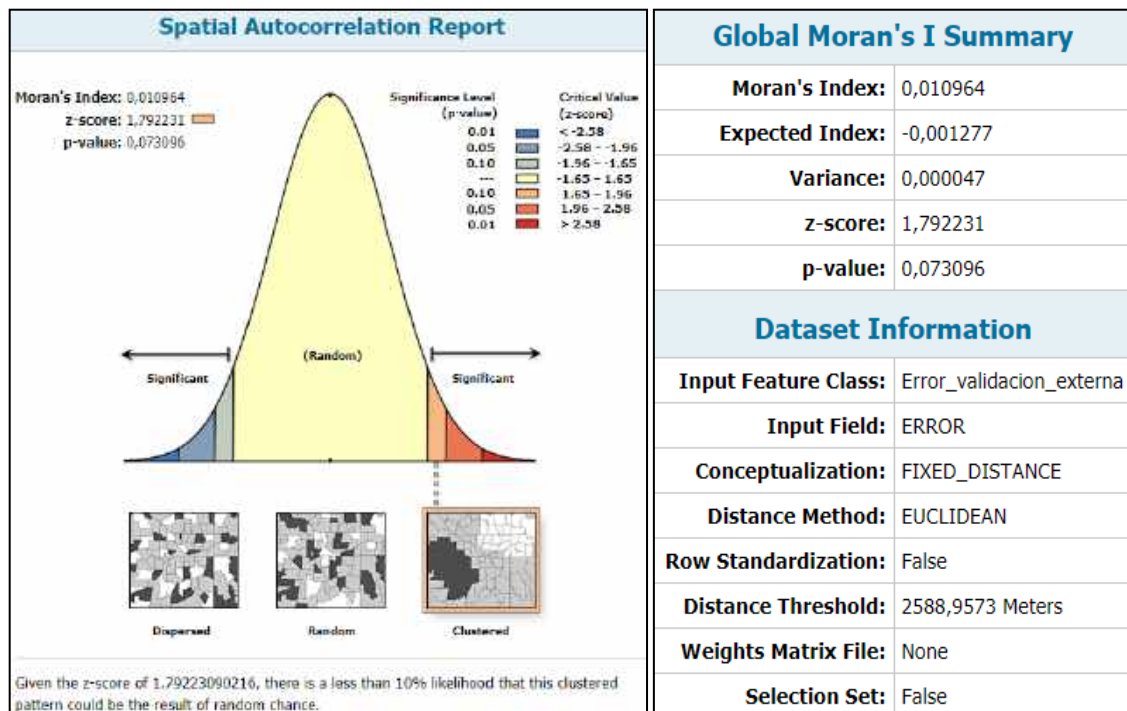


Figura 44: Índice de Moran para validación cruzada externa (modelo R)

4.6 ESTABILIDAD DE LOS MODELOS

Como se mencionó en el apartado correspondiente, el objetivo de utilizar el índice de Moran fue para establecer el tipo de tendencia seguida por los errores existentes sobre los modelos generados, y, así determinar cuál de éstos era el más estable. En la Tabla 8 se presenta un resumen de los resultados obtenidos al realizar dicho análisis. Los valores de la Tabla permiten apreciar varias discrepancias en función a las dos validaciones cruzadas, que se han observado y aplicado en este TFM.

Tabla 8: Resumen de estabilidad de residuales en modelos analizados

| RESUMEN ESTABILIDAD (MORAN) | | | | |
|------------------------------------|---------------------------------------|-----------------------|---------------------------------------|-------------------------------------|
| MODELOS DE ESTIMACIÓN | Kriging ArcGIS | | Kriging R-Studio | |
| Método evaluación | VCI (1) | VCE (2) | VCI | VCE |
| Z | -5,80 | -1,22 | -6,15 | 1,79 |
| P (%) | < 1 | 22* | < 1 | 7* |
| Observaciones | Heterocedasticidad en Zonas dispersas | No Heterocedasticidad | Heterocedasticidad en Zonas dispersas | Heterocedasticidad en Zonas clúster |

* Significancia estadística, no representativa

(1) Validación cruzada interna

(2) Validación cruzada externa

Cuando se analizaron los resultados del índice de Moran, aplicado a la validación cruzada interna de ambos modelos kriging, se comprobó que presentan heterocedasticidad en puntos dispersos sobre sus respectivas superficies de estimación. Estos puntos, que se han identificado a pesar de ser independientes y se han representado en las Figuras 45 y 46 y en los Anexos 8 y 9, presentan patrones de distribución similares entre sí, con un nivel de confianza de más del 99 % de sufrir heterocedasticidad.

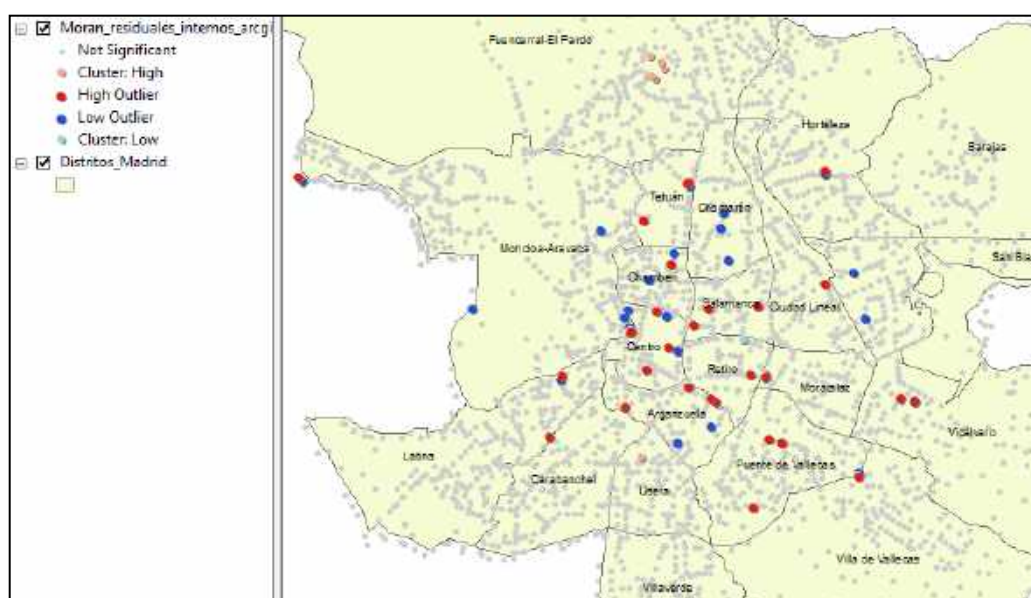


Figura 45: Zonas con heteroscedasticidad sobre superficie de estimación kriging generada con ArcGIS, para validación cruzada interna (ver mapa detallado en el Anexo 8)

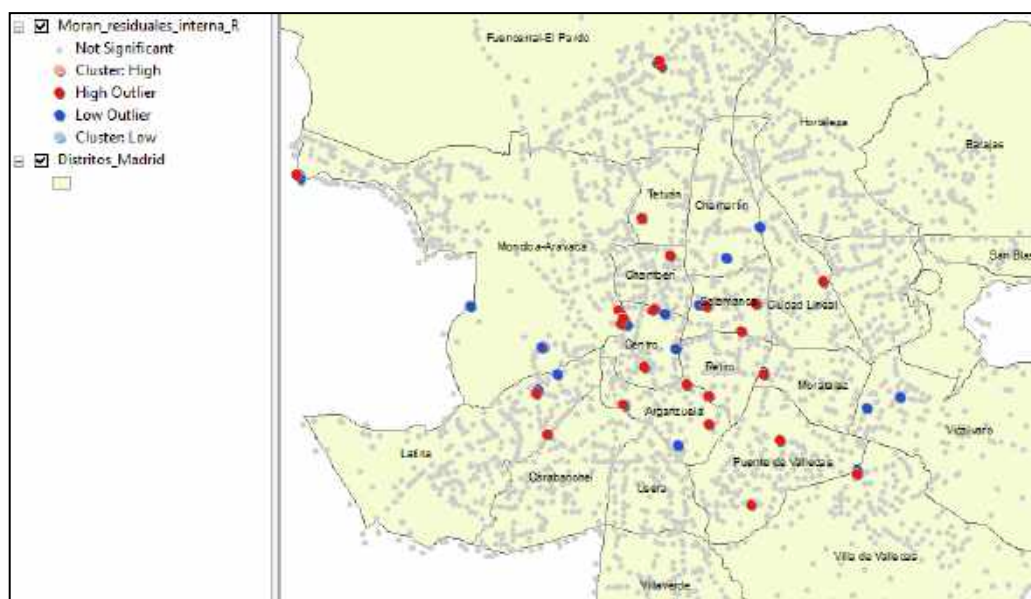


Figura 46: Zonas con heteroscedasticidad sobre superficie de estimación kriging generada con R, para validación cruzada interna (ver mapa detallado en el Anexo 9)

Por otra parte, los resultados del índice de Moran, aplicado a la validación cruzada externa (expuestos en la tabla 8), denotan comportamientos diferentes para cada modelo. Así, en el modelo kriging generado con ArcGIS, el índice de Moran determina que no existe heteroscedasticidad significativa, tal como se manifiesta en la figura 47 y Anexo 10. El mapa representa la distribución de posibles zonas aleatorias con un nivel de confianza del 78% de exhibir este fenómeno. No obstante, para el modelo kriging generado con R, también se determinó que no existe heteroscedasticidad significativa; empero, para este modelo existe mayor probabilidad de encontrar zonas clusters con heteroscedasticidad (Figura 48 y Anexo 11) con niveles de confianza del 93%.

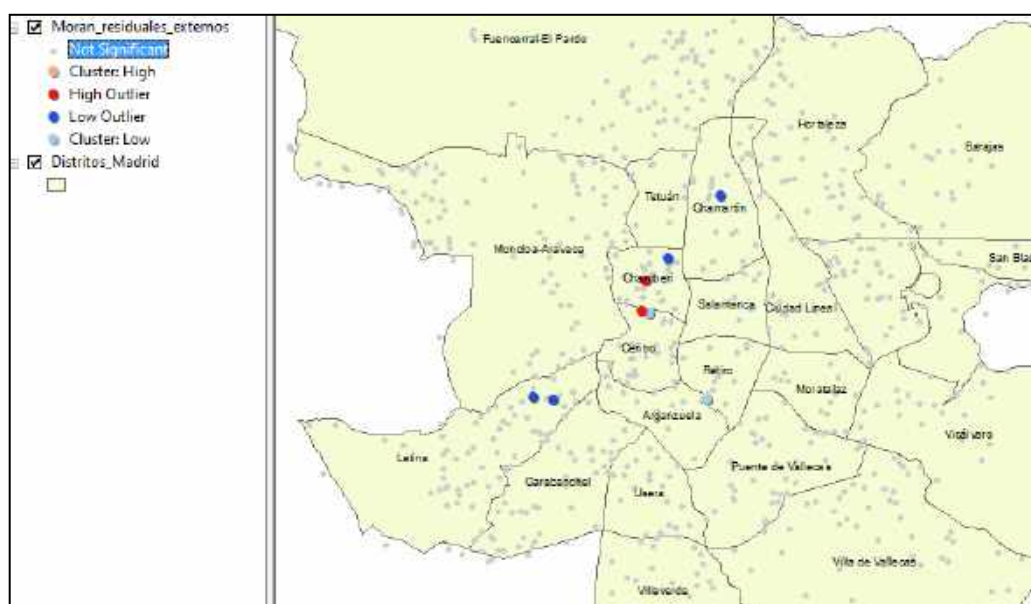


Figura 47: Zonas con heteroscedasticidad sobre superficie de estimación kriging generada con ArcGIS, para validación cruzada externa (ver mapa detallado en el Anexo 10)

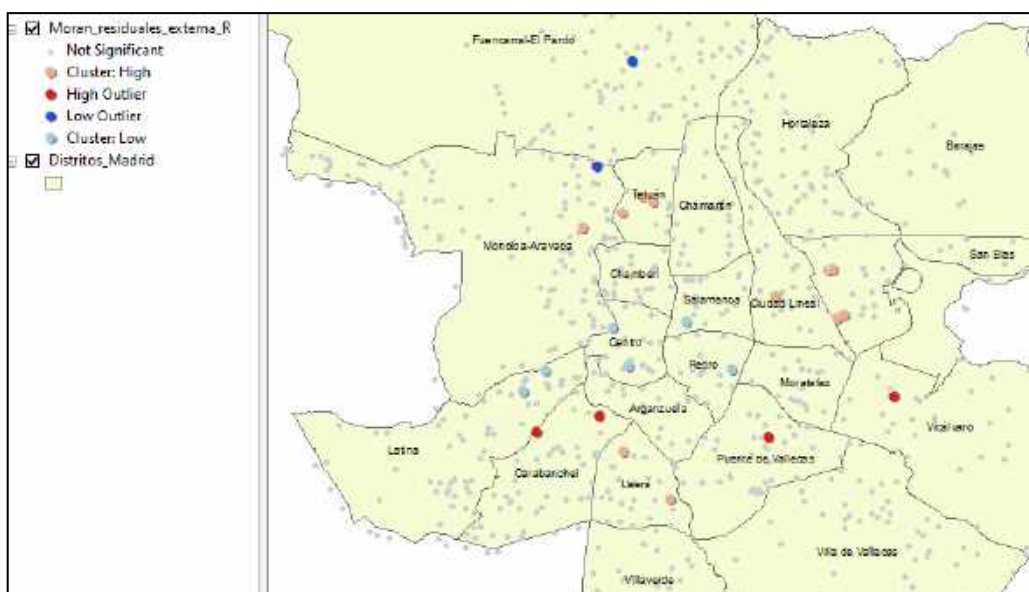


Figura 48: Zonas con heteroscedasticidad sobre superficie de estimación kriging generada con R, para validación cruzada externa (ver mapa detallado en el Anexo 11)

Por lo tanto, debido a que ambas validaciones ofrecen criterios opuestos, se tomó como verdadero el resultado que ofrece la validación externa. Esto último responde a los criterios sugeridos por Cañadas (2013) y Giraldo (2007), quienes mencionan que la validación cruzada externa debe ser utilizada para determinar la calidad de un modelo, descartando la observación de un grupo (representativo) de los “n” puntos muestrales escogidos para modelar al variograma, dado que el objetivo, es predecir mediante Kriging el valor de la variable en dichos puntos excluidos. Por esta razón, se determina que el modelo con mayor estabilidad para estimar el valor de N sobre la ciudad de Madrid, es el modelo kriging generado por ArcGIS.

Finalmente, debido a las grandes diferencias encontradas al comparar los resultados de cada validación cruzada, se analizó aquellos puntos de la validación interna que tienen la mayor probabilidad de generar heteroscedasticidad en sus respectivos modelos. En la Tabla 9, se observa los distritos que poseen puntos que afectan a la estabilidad de cada modelo, en especial los vértices que poseen residuales con una baja correlación entre sí (correlaciones HL y LH). Por lo tanto, a estos puntos se los puede considerar como datos atípicos de la base del RTM, ya que al sumar los mismos con los 37 puntos depurados en el apartado 3.3, no llegan ni al 3% de la población total.

Tabla 9: Vértices con residuales altos en función a su entorno, localizados por Moran.

| DISTRITO | KRIGING ARCGIS | | | | KRIGING R | | | |
|---------------------|-------------------|------------------|------------------|-----------------|-------------------|------------------|------------------|-----------------|
| | HH (High-High) | HL (High-Low) | LH (Low-High) | LL (Low-Low) | HH (High-High) | HL (High-Low) | LH (Low-High) | LL (Low-Low) |
| Centro | 1 | 6 | 3 | | 1 | 3 | 3 | |
| Fuencarral-El Pardo | 6 | | | | | 2 | 1 | |
| Usera | 1 | | | | | | | |
| Arganzuela | | 2 | 3 | | | 3 | 4 | |
| Carabanchel | | 1 | | | | 1 | | |
| Chamberí | | 1 | 2 | | | 1 | 2 | |
| Ciudad Lineal | | 1 | | | | 1 | 1 | |
| Hortaleza | | 1 | 1 | | | | | |
| Latina | | 1 | 2 | | | 1 | 4 | |
| Moncloa-Aravaca | | 1 | 5 | | 1 | 3 | 3 | |
| Moratalaz | | 1 | 1 | | | 1 | 1 | |
| Puente de Vallecas | | 3 | 2 | | | 2 | 2 | |
| Retiro | | 3 | | 1 | | 1 | | |
| Salamanca | | 3 | | | | 3 | 2 | |
| Tetuán | | 3 | 1 | | | 1 | | |
| Vicálvaro | | 2 | 1 | | | | 2 | |
| Villa de Vallecas | | 1 | | | | 1 | | |
| Chamartín | | | 3 | | | | | |
| San Blas | | | 2 | | | | | |
| TOTAL | 8 | 30 | 26 | 1 | 2 | 24 | 25 | 0 |

Por otra parte, se podría considerar que los vértices con correlaciones HL y LH, pueden tener relación con aquellas zonas con cambios bruscos en donde el valor de N es más elevado debido a la topografía de la ciudad. Sin embargo, si se coteja dichos vértices con los mapas de los Anexos 4 y 6, se observa que no existe relación alguna ya que los distritos con alturas elevadas como Tetuán, Ciudad Lineal, Moratalaz, San Blas y Carabanchel, son levemente afectados por vértices con correlación HL y LH. Mientras que los distritos que si poseen diversos vértices HL y LH son Moncloa-Aravaca, Puente de Vallecas, Arganzuela, Centro, Salamanca y la Latina, pero estos no presentan cambios bruscos en los valores de N a lo largo de toda la zona de estudio.

CAPITULO V. CONCLUSIONES Y RECOMENDACIONES

5.1 CONCLUSIONES

- En este Trabajo Fin de Máster (TFM) se obtuvieron 2 modelos de predicción de la variable, ondulación geoidal (N) para la ciudad de Madrid, utilizando técnicas geoestadísticas a través de software libre (R-Studio) y comercial (ArcGIS). La calidad de estos modelos fue valorada por medio de la validación cruzada, alcanzando precisiones menores a 1.2 cm, con un nivel de confianza del 95%. Por lo tanto, cualquiera de estos dos modelos podrá ser utilizado para cualquier aplicación de ingeniería que requiera dicha precisión en el área de estudio.
- Al comparar el RMSE de los 2 modelos generados se observó una diferencia inferior al milímetro entre ellos, siendo el modelo y la superficie de estimación creada en ArcGIS más precisa. Sin embargo, a pesar de lo afirmado, R-Studio permitió una mayor libertad en el momento de realizar el análisis de estructura del semivariograma, así como una mayor facilidad al ejecutar diferentes pruebas para el ajuste del mismo. Por lo tanto, se concluye que las diferencias entre software libre y comercial, obtenidas en este trabajo fueron mínimas.
- Por medio del índice de autocorrelación espacial de Moran fue posible determinar el grado de estabilidad que posee cada modelo local, al momento de estimar el valor de N. Dicho índice, fue aplicado a las 2 validaciones cruzadas estudiadas, obteniendo diversos criterios debido a la significancia estadística de cada prueba. No obstante, en referencia a los estudios de Giraldo (2007) y Cañadas (2013), se dio mayor peso a la validación externa en el momento de verificar la estabilidad de los modelos. Se concluye que la superficie de estimación conseguida con ArcGIS es aquella que posee residuales con una distribución normal a lo largo de área de estudio.
- Con el objetivo de dar una mayor validez a este TFM, se cotejaron los modelos locales geoestadísticos con el modelo español EGM08-REDNAP del IGN, en los mismos puntos usados para la validación cruzada externa, y se consideró el RMSE como criterio de evaluación. Los resultados revelan que los modelos locales se ajustan mucho mejor y son 25 veces más precisos que el modelo EGM08-REDNAP para el municipio de Madrid. Sin embargo, esta precisión se debe a dos factores: el primero a la altísima densidad de vértices o puntos que posee la RTM con respecto a la REDNAP (aproximadamente 20 veces más densa). Mientras que el segundo se relaciona con el método de interpolación utilizado en este TFM, el cual es un método probabilístico, mientras que el modelo REDNAP utilizó uno determinístico. Por todo lo mencionado, se concluye que los resultados de un modelo geoestadístico dependen directamente de su información base, así como el tipo de estimador utilizado. En este sentido, el fenómeno de la ondulación Geoidal es definido por una variable regionalizada (variable espacialmente correlacionada), la cual no debe ser modelada por métodos determinísticos, debido a su extrema complejidad y tampoco considerar que estas observaciones son variables aleatorias independientes a su posición.

5.2 RECOMENDACIONES

- Si se analiza el índice de autocorrelación espacial de Moran para la validación cruzada interna, se observa que existe cierto grado de heteroscedasticidad en ambos modelos. Por ende, se recomienda descartar aquellos vértices que poseen residuales con el comportamiento de HL (High Outlier) y LH (Low Outlier), tal como se mostraba en la tabla 9. Estos residuales pueden considerarse como datos atípicos al no superar ni el 3% de población muestral. Al descartar estos residuales los modelos locales podrían mejorar su estabilidad de predicción.
- Se recomienda utilizar para aplicaciones de ingeniería, topografía o construcción, dentro de los límites del municipio de Madrid, cualquiera de los 2 modelos de estimación Kriging obtenidos en este trabajo, en vez del modelo EGM08-REDNAP, debido a las precisiones milimétricas que estos lograron alcanzar al estimar el valor de la ondulación geoidal.
- Se recomienda el uso de software libre, puesto que es posible obtener resultados muy similares a los proporcionados por software comercial, abaratando costos y adquiriendo libertad de uso. Sin embargo, se debe tener en cuenta que el usuario de software libre necesita tener otro tipo de capacitación para la manipulación del código abierto, ya que puede generar productos cuestionables al no poseer controles de calidad, ni asistencia al usuario como lo haría un software licenciado.
- Para otros proyectos similares a este, que modelan variables espacialmente correlacionadas, en donde sus resultados son productos que pueden ser utilizados por otros usuarios, se recomienda, realizar diferentes pruebas del fenómeno que se desea modelar, así como, a ser posible, el uso conjunto de software libre y comercial, ya que el producto final gozara de las ventajas que cada uno le puede brindar.

BIBLIOGRAFÍA

- Alfaro Sironvalle, M. (2007). *Estimación de recursos mineros*. Obtenido de Universidad de Católica de Valparaíso: http://cg.ensmp.fr/bibliotheque/cgi-bin/public/bibli_index.cgi
- Arce, R. (2001). Conceptos básicos sobre la heterocedasticidad en el modelo básico de regresión lineal tratamiento con E-VIEWS. Dpto. de Economía Aplicada, Universidad Autónoma de Madrid. Madrid, España.
- Ayuntamiento de Madrid. (2019). *Ayuntamiento de Madrid*. Obtenido de datos abiertos: <https://datos.madrid.es/portal/site/egob/menuitem.c05c1f754a33a9fbe4b2e4b284f1a5a0/?vgnextoid=49713c7cfc981610VgnVCM1000001d4a900aRCRDyvgnextchannel=374512b9ace9f310VgnVCM100000171f5a0aRCRDyvgnextfmt=default>
- Barahona, C. (2016). Metodología para la determinación del cuasigeoide para el Ecuador continental aplicando la teoría de Molodensky. Quito, Ecuador: Facultad de Ciencias de la Tierra y Construcción - ESPE. Obtenido de Monografías e informes: <https://repositorio.espe.edu.ec/bitstream/21000/12115/1/T-ESPE-053424.pdf>
- Blitzkow, D., Drewes, H., Sánchez, L., y Freiras, S. (S/F). *Sistema vertical de referencia para América del Sur*. Obtenido de <http://azimuth.univalle.edu.co/docsdownload/archivo3.doc>
- Bosque Sendra, J. (2000). *Sistemas de Información Geográfica*. Madrid, España: Ediciones Rialp.
- Buill Pozuelo, A. F., y Núñez, M. A. (2005). Necesidad del geoide en el posicionamiento con GPS. Departamento de Ingeniería del Terreno, Cartografía y Geofísica. Universidad Politécnica de Cataluña. Barcelona, España.
- Cañadas, J. (2013). *Regresión logística y tratamiento computacional con R*. Facultad de Ciencias. Universidad de Granada. Granada, España. Obtenido de https://masteres.ugr.es/moea/pages/tfm-1213/tfm_caaadasreche_jluis/
- Chica Olmo, M., y Luque Espinar, J. A. (2002). Interpolación espacial en la creación de cubiertas temáticas en S.I.G. En: L. Lain Huerta (Eds). *Los Sistemas de Información Geográfica en la Gestión de los Riesgos Geológicos y el Medio Ambiente*. Publicaciones del Instituto Geológico y Minero de España. Serie Medio Ambiente. Riesgos Geológicos, nº 3: 181-198.
- Clark, I. (1979). *Practical Geostatistics*. Elsevier Publishing, New York, USA.
- Desassis, N., y Renard, D. (2013). Automatic variogram modeling by iterative least squares: Univariate and multivariate cases. *Math Geosci*, 2013 (45), 453- 470. doi: 10.1007/s11004-012-9434-1.
- Drewes, H. (2014). *Sistemas de Referencia*. Instituto Geográfico Militar. Quito, Ecuador.
- Emery, X. (2013). *Geoestadística*. Facultad de ciencias Físicas y Matemáticas. Universidad de Chile. Santiago de Chile, Chile.

- ESRI. (2019). *Esri for Desktop*. Obtenido de Tools:
<http://desktop.ArcGIS.com/en/arcmap/10.3/main/tools/a-quick-tour-of-geoprocessing-tool-references.htm>
- Fernández, I. (2012). Localizaciones Geográficas. *El Datum*. Area de Ingeniería Cartográfica, Geodesia y Fotogrametría. Universidad de Valladolid. Valladolid, España.
- Freitas, S., y Blitzkow, D. (1999). Altitudes y geopotencial. International Geoid Service Bulletin No. 9: p. 47-62. Milan, Italia.
- Furones, A. (2011). *Sistema y marco de referencia terrestre*. Universidad Politécnica de Valencia. Valencia, España
- Furones, A., Padin, J., y Garcia, F. (2000). *Apuntes de Geodesia Física*. Servicio de publicaciones UPV. spupv-2000.869. Valencia, España. Obtenido de <http://personales.upv.es/jpadin/TEMA%206.pdf>
- Giraldo, R. (2007). Introducción a la Geoestadística. *Teoría y Aplicación*. Departamento de Estadística, Universidad Nacional de Colombia. Bogota, Colombia.
- Goovaerts, P. (1997). *Geostatistics for natural resources valuation evaluation*. Oxford University Press. New York, USA.
- Heiskanen, W., y Moritz, H. (1985). Geodesia Física (Primera ed). Instituto Geográfico Nacional de España. Madrid, España.
- Hernández, D. (S/F). Geodesia y Cartografía Matemática. Escuela Técnica Superior de Ingeniería Geodésica, Cartográfica y Topográfica. Universidad Politécnica de Valencia. Valencia, España
- Instituto Geográfico Nacional. (2009). Nuevo modelo de Geoide para España EGM08-REDNAP. Centro de Observaciones Geodésicas, Subdirección General de Astronomía, Geodesia y Geofísica. Madrid, España.
- Isaaks, E., y Srivastava, R. M. (1989). *Applied Geostatistics*. Oxford University. Oxfordshire, England
- Leick, A. (2004). *GPS Satellite Surveying*. Jonh Wiley y Sons, Inc. New jersey, USA.
- Leiva, C. (2014). Determinación de modelos de predicción espacial de la variable ondulación geoidal, para la zona urbana del Cantón Quito y la zona rural del Cantón Guayaquil, utilizando técnicas geoestadísticas. Facultad de Ciencias - Escuela Politecnica Nacional. Quito, Ecuador.
- Maune, D. F., Kopp, S. M., Crawford, C. A., y Zervas, C. E. (2001). *Introduction digital elevation model technologies and applications: the DEM users manual*. The American Society for Photogrammetry and Remote Sensing. Maryland, USA.
- Moirano, J. (2000). Materialización del sistema de referencia terrestre Internacional en Argentina mediante observaciones GPS. Universidad Nacional de La Plata. La Plata, Argentina.
- Moritz, H. (1984). *Sistemas de referencia en geodesia*. Instituto de Astronomía y Geodesia. Madrid, España

- Oliver, M., y Webster, R. (2010). Basic steps in geostatistics: The Variogram and Kriging. Springer Science+Business Media BV. SpringerBriefs in Agriculture. New York.
- Paredes, C., Salinas, W., Martinez, X., y Jimenez, S. (2012). *Evaluación y comparación de métodos de interpolación determinísticos y probabilísticos para la generación de modelos digitales de elevación*, 3 y 4. Num 82: Investigaciones Geográficas, Boletín del Instituto de Geografía, UNAM.
- Sánchez, L. (2005). Definition and realisation of the SIRGAS vertical Reference System within a Global Unified Height System. *De P. Tregoning, y C. Rizos, Dynamic Planet. Monitoring and understanding a dynamic planet with Geodetic and Oceanographic tools*. International Association of Geodesy Symposia.
- Seeber, G. (2003). Satellite Geodesy. Walter de Gruyter. New York, USA.
- Torge, W. (2001). Geodesy. Walter de Gruyter. New York, USA.
- Zakatov, P. S. (1997). Curso de Geodesia Superior. Rubiños-1860. Madrid, España.
- Zhang, S. (2000). Interpolation of Geoidal/Quasigeoidal surfaces for height determination with GPS. Wiss. Arb. Univ. Hannover, Nr. 236.

ANEXOS

ANEXO 1: SCRIPT Generado EN R-STUDIO PARA MODELAR LA ONDULACIÓN GEOIDAL DE MADRID

CARGAR LIBRERIAS PARA ANALISIS GEOESTADISTICOS

```
library(Rcpp)
library(RGeostats)
library(Rcmdr)
library(Rcpp)
library(ggplot2)
library(lattice)
library(caret)
```

#PARA CARGAR BASE DE DATOS

```
setwd("D:/Rai/Documents/Master TICS UCM/TFM/Ondulacion geoidal/Restudios/Datos
depurados")
getwd()
bd<-read.csv("Datos_depurados_N1.csv")
head(bd) #Visualizar datos en tabla
help("hist")
```

#PARTICION DE BASE DE DATOS

```
data.ids<-createDataPartition(bd$ID,p=0.8,list=FALSE) # PARTIR LA BASE AL 80%
DE DATOS
bd_p<-bd[data.ids,] #CAMBIO DE VARIABLE A LA BASE PRINCIPAL DIVIDIDA
bd_v<-bd[-data.ids,] #CAMBIO DE VARIABLE A LA BASE PARA LA VALIDACION
(20% DE DATOS)
write.csv(bd_v,file = "validation_externa_data.csv") #EXPORTA EN FORMATO CSV
DATOS PARA LA VERIFICACION DE VALIDACION CRUZADA EXTERNA
```

#ANALISIS EXPLORATORIO DE DATOS

```
summary(bd_p) #ESTADISTICOS GENERALES DE CADA VARIABLE
hist(bd_p$ondulacion, col = "blue") #DEVUELVE HISTOGRAMA
boxplot(bd_p$ondulacion, add = FALSE, col = "blue",plot = TRUE,) #DEVUELVE
DIAGRAMA DE CAJA
mean(bd_p$ondulacion) #MEDIA
sd(bd_p$ondulacion) #DESVIACION ESTANDAR
cv<-sd(bd_p$ondulacion)/mean(bd_p$ondulacion) # CALCULAR EL COEFICIENTE DE
VARIACION
cv #VISUALIZAR EL COEFICIENTE DE VARIACION, EL CUAL DEBERIA SER
<50%
```

#PARA IDENTIFICAR SI EL PROCESO ES ESTACIONARIO

```
Boxplot(~ondulacion,data=bd_p,id.method="y") #DEVUELVE DIAGRAMA DE CAJA
CON OTRA FUNCION
scatterplot(ondulacion~X_ETRS89,reg.line="lm",smooth=TRUE,spread=TRUE,id.method
='mahal',id.n=2,boxplots="xy",span=0.5,data=bd) #PARA GRAFICAR LA DISPERSION
DATOS PARA PROCESO ESTACIONARIO EN "X"
scatterplot(ondulacion~Y_ETRS89,reg.line="lm",smooth=TRUE,spread=TRUE,id.method
='mahal',id.n=2,boxplots="xy",span=0.5,data=bd) #PARA GRAFICAR LA DISPERSION
DATOS PARA PROCESO ESTACIONARIO EN "Y"
```

```

#GENERACION DE OBJETO EN GEOSTATS
Ndb<-db.create(x1=bd_p$X_ETRS89,x2=bd_p$Y_ETRS89,z1=bd_p$ondulacion)
#CREAR BASE DE DATOS
plot(Ndb,pch=21,bg="blue",col="black",tittle="samples location") #VIZUALIZACION
ESPACIAL DE MUESTRAS
data.4dir.vario=vario.calc(Ndb,dir=c(0,45,90,135),lag = 500,tolang=22.5)
#GENERACION DE VARIOGRAMA DIRECCIONALES, TOLANG ES EL ANGULO
DE BARRIDO/2
data.4dir.vario #PARA VISIALIZAR LA GENERACION DEL VARIOGRAMA
plot(data.4dir.vario,title="Direccional Variograma") #PARA GRAFICAR
VARIOGRAMA

#PARA PLOTEAR UNICAMENTE VARIOGRAMAS EN UN DIRECCION
ESPECIFICA
plot(data.4dir.vario,idir0=1,title="Directionalvariogram0(BLUE) -
Directionalvariogram90(GREEN)",npairdw=TRUE,npairpt=1,col="BLUE")
#VARIOGRAMA 0° AZUL
plot(data.4dir.vario,idir0=2,title="Directionalvariogram45",npairdw=TRUE,npairpt=1,add
=TRUE,col="RED") #VARIOGRAMA 45° ROJO
plot(data.4dir.vario,idir0=3,title="Directionalvariogram90",npairdw=TRUE,npairpt=1,add
=TRUE,col="GREEN") #VARIOGRAMA 90° VERDE
plot(data.4dir.vario,idir0=4,title="Directionalvariogram135",npairdw=TRUE,npairpt=1,ad
d=TRUE,col="BLACK") #VARIOGRAMA 135° NEGRO

#GENERACION DE VARIOGRAMA EXPERIMENTAL OMNIDIRECCIONAL
vario_data<-vario.calc(Ndb, lag=300, tolang = 22.5, bench = 3, calcul="vg")
vario_data #VISUALIZACION DE VARIOGRAMA OMNIDIRECCIONAL
vario.plot(x=vario_data,npairdw=TRUE,add=FALSE,npairpt=1,title="Variograma
experimental Omnidireccional") #GRAFICAR VARIOGRAMA OMNIDIRECCIONAL

#AJUSTE DE MODELO PARA VARIOGRAMA TEORICO
data_model<-model.auto(vario_data,struct=melem.name(c(1,2,4,11)), auth.aniso =
TRUE,col=2) #AJUSTE DE MODELO TEORICO, LISTA DE FUNCIONES EN
ANEXO 2
str(data_model) #VISUALIZAR LOS PARAMETROS DEL MODELO OPTIMIZADO

#NUBE VARIOGRAFICA
nube_va<-cloud.calc(Ndb,dbout = NA,dirvect = NA,lagmax = 22000,varmax = 4,lagnb =
100,varnb = 100,toldis = 0.5,tolang = NA,opt.code = 0,tolcode = 0) #generacion de nube
db.plot(nube_va,pos.x = 1,pos.y = 2,title = "Nube Variografica",add=FALSE,reset =
TRUE, ylim =c(0,0.04),xlim = NA) #VISUALIAR NUBE

#VECINDAD DE ESTIMACION (MATRIZ DE BUSQUEDA)
moving.neigh<-neigh.input(ndim = 2) #POR VECINOS MOVILES CERCANOS

#VALIDACION CRUZADA INTERNA
data.db1<-xvalid(Ndb,data_model,moving.neigh) #GENERAR VALIDACION
CRUZADA INTERNA
data.db1 #VISUALIZAR

```

```

hist(db.extract(data.db1,names=c("Xvalid.z1.esterr")),nclass = 50,main = "ERROR DE
VALIDACION CRUZADA",xlab = "Error validacion cruzada",col="BLUE")
#VISUALIZAR HISTOGRAMA DE ERRORES ESTIMADOS
db.write(data.db1,filename="errorvalidacion_interna.csv",must.noproj = TRUE,nogrid =
TRUE,flag.calcul = TRUE) #EXPORTAR BASE EN FORMATO CSV (ERRORES DE
PREDICCION)

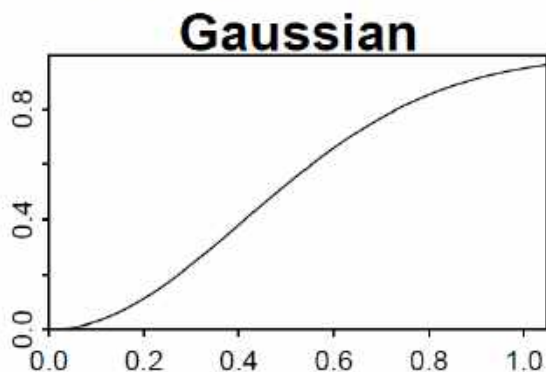
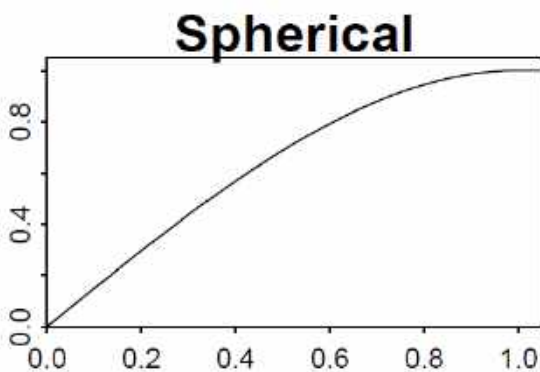
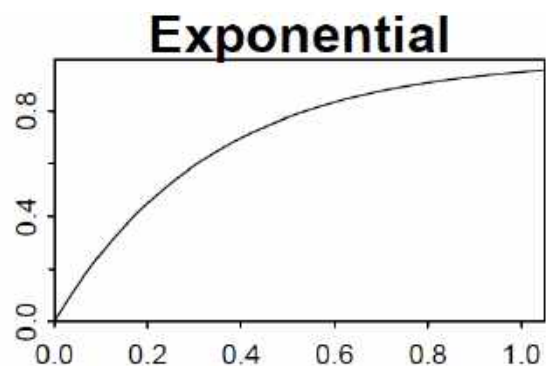
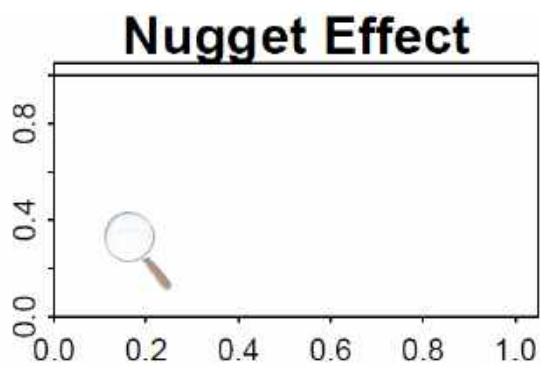
#GENERACION DE GRILLA PARA KRIGIN
griddb<-db.create(flag.grid =
TRUE,nx=c(540,553),x0=c(428939,4462566),dx=c(50,50),autoname = FALSE)
#FORMAR GRILLA EN FUNCION A COLUMNAS, FILAS Y RESOLUCION
ESPACIAL
print(griddb) # VISUALIZAR DATOS DE GRILLA

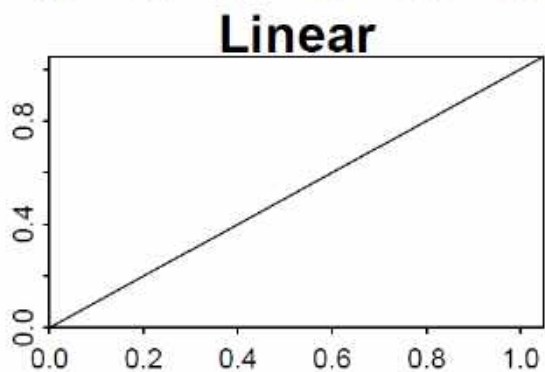
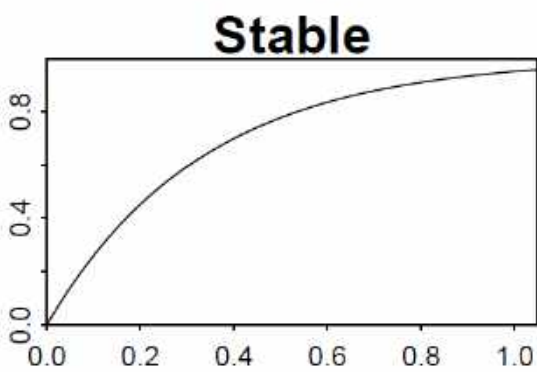
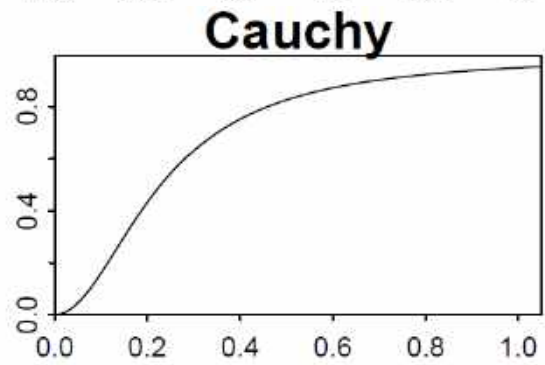
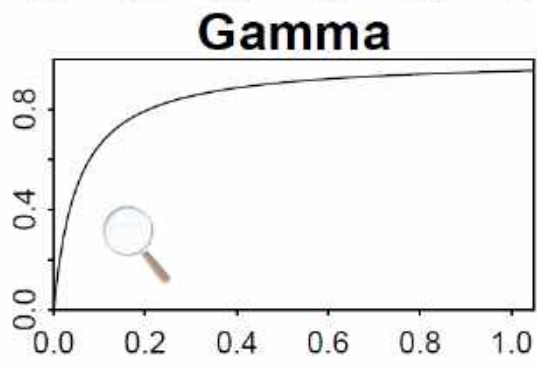
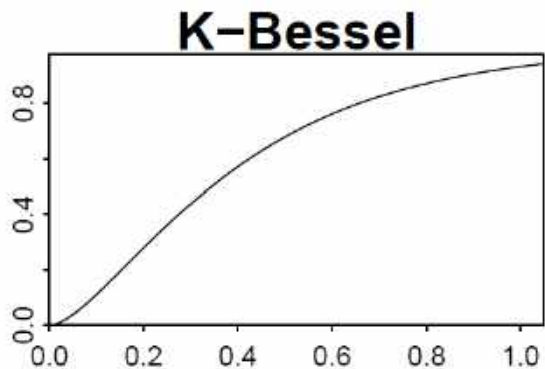
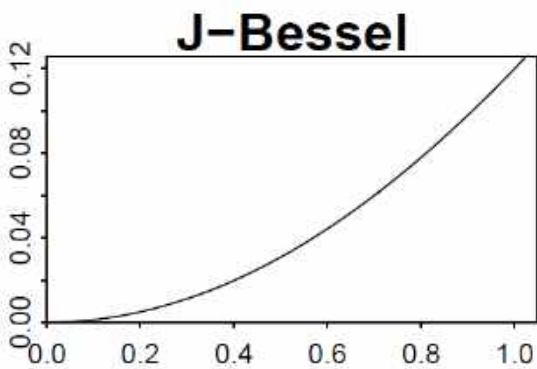
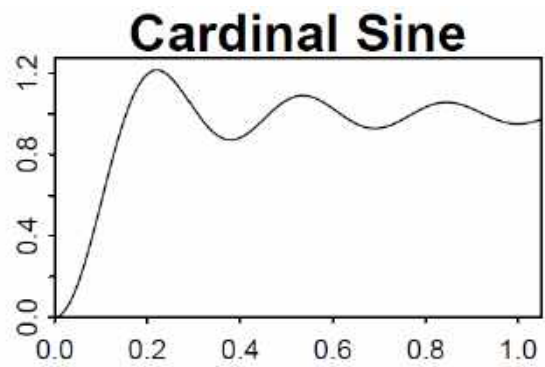
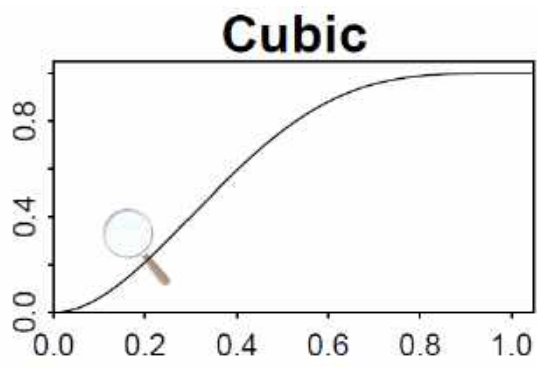
#KRIGIN
griddb<-kriging(Ndb,griddb,data_model,moving.neigh,radix="N.ALL") #MODELO
KRIGIN
griddb
plot(griddb,name.image="N.ALL.z1.estim",col=topo.colors(40),title="Estimacion -All
data(Moving Neighborhood)") #Visualizar KRIGING como raster
plot(griddb,name.contour="N.ALL.z1.estim",nlevels=20,add=TRUE) #VISUALIZAR
KRIGING COMO RASTER CON ISOLINEAS
db.write(griddb,filename = "modelo_N_prediccion_krigin.asc",must.noproj =
TRUE,nogrid = TRUE,flag.calcul = TRUE) # EXPORTAR FORMATO ASC GRILLA
CON KRIGING
db.write(griddb,filename = "modelo_N_prediccion_krigin.csv",must.noproj =
TRUE,nogrid = TRUE,flag.calcul = TRUE) # EXPORTAR FORMATO CSV GRILLA
CON KRIGING

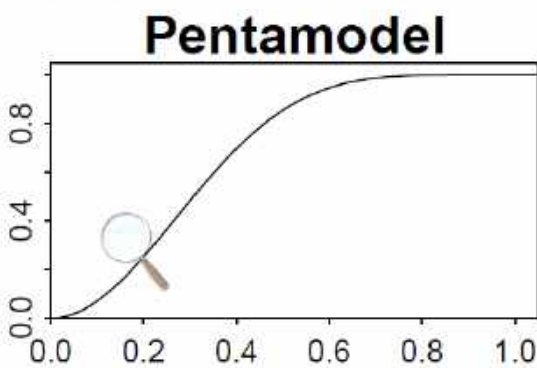
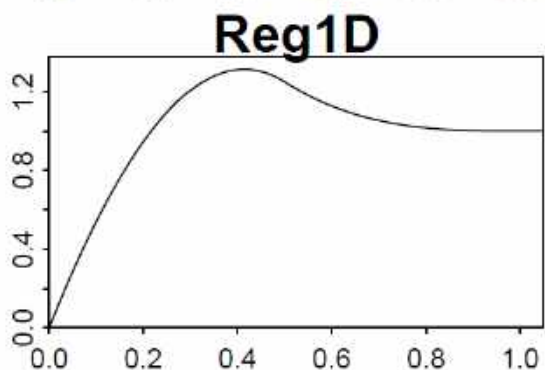
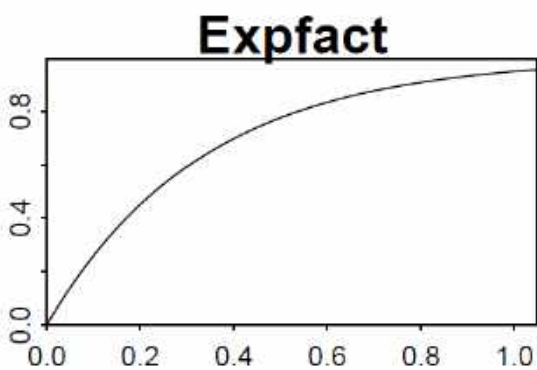
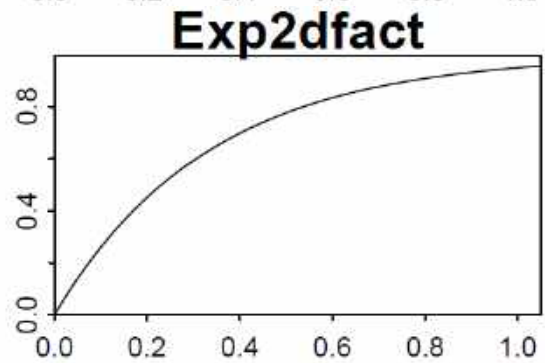
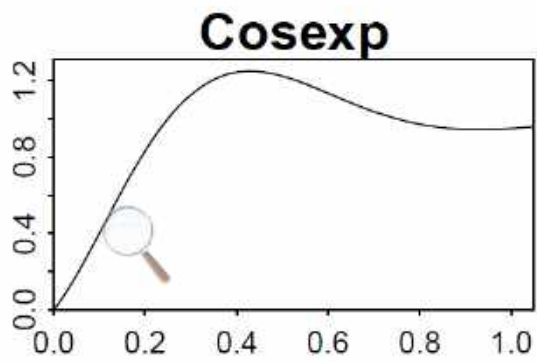
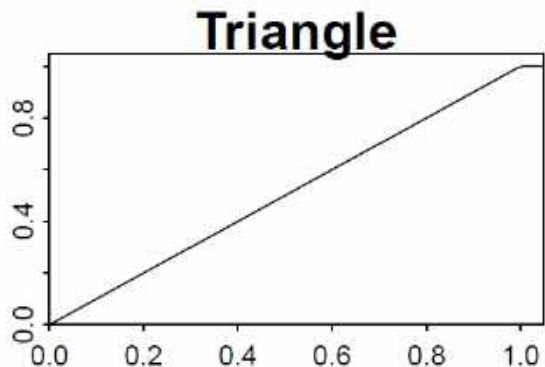
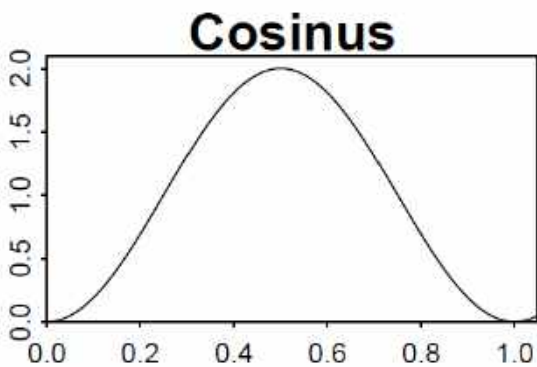
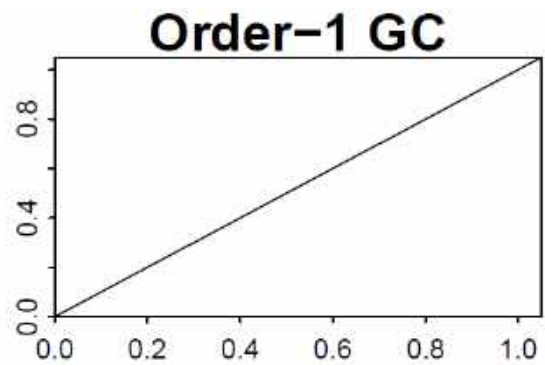
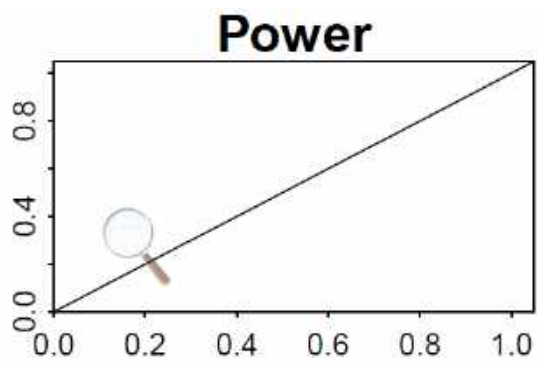
```


ANEXO 2: MODELOS TEÓRICOS PARA VARIOGRAMA - PAQUETE RGEOSTAT – R
Recuperado de Leiva (2014).

- 1 - Nugget Effect
- 2 - Exponential
- 3 - Spherical
- 4 - Gaussian
- 5 - Cubic
- 6 - Cardinal Sine
- 7 - J-Bessel
- 8 - K-Bessel
- 9 - Gamma
- 10 - Cauchy
- 11 - Stable
- 12 - Linear
- 13 - Power
- 14 - Order-1 GC
- 15 - Cosexp
- 16 - Exp2dfact
- 17 - Expfact







ANEXO 3: RESUMEN DE CONSOLA EJECUTADO EL SCRIPT EN R-STUDIO

R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> # CARGAR LIBRERIAS PARA ANALISIS GEOESTADISTICOS

> library(Rcpp)

Warning message:

package 'Rcpp' was built under R version 3.5.3

> library(RGeostats)

Loading RGeostats - Version:11.2.3

> library(Rcmdr)

Loading required package: splines

Loading required package: RcmdrMisc

Loading required package: car

Loading required package: carData

Loading required package: sandwich

Error: package or namespace load failed for 'RcmdrMisc' in loadNamespace(j <- i[[1L]], c
(lib.loc, .libPaths()), versionCheck = vI[[j]]):

there is no package called 'latticeExtra'

Error: package 'RcmdrMisc' could not be loaded

In addition: Warning messages:

1: package 'Rcmdr' was built under R version 3.5.3

2: package 'RcmdrMisc' was built under R version 3.5.3

3: package 'car' was built under R version 3.5.3

4: package 'carData' was built under R version 3.5.3

5: package 'sandwich' was built under R version 3.5.3

> library(Rcpp)

> library(ggplot2)

Warning message:

package 'ggplot2' was built under R version 3.5.3

> library(lattice)

Warning message:

package 'lattice' was built under R version 3.5.3

> library(caret)

Warning message:

package 'caret' was built under R version 3.5.3

```

> #PARA CARGAR BASE DE DATOS
> setwd("D:/Rai/Documents/Master TICS UCM/TFM/Ondulacion geoidal/Restudios/Datos depurados")
> getwd()
[1] "D:/Rai/Documents/Master TICS UCM/TFM/Ondulacion geoidal/Restudios/Datos depurados"
> bd<-read.csv("Datos_depurados_N1.csv")
> head(bd) #Visualizar datos en tabla
  ID Vertice Hoja_MTN  X_ED50  Y_ED50 X_ETRS89 Y_ETRS89 Z_Ortometr Z_elipsoidal ondulacion
1 1  1339    559 438249.4 4479279 438140.1 4479072  655.360   706.379   51.019
2 2  1340    559 438296.0 4479396 438186.6 4479189  653.852   704.871   51.019
3 3  1341    559 438273.6 4479519 438164.3 4479311  654.156   705.176   51.020
4 4  1342    559 438282.8 4479632 438173.5 4479425  657.473   708.493   51.020
5 5  1343    559 438156.8 4479658 438047.4 4479450  663.971   714.991   51.020
6 6  1344    559 438130.5 4479877 438021.1 4479669  670.933   721.953   51.020
> help("hist")
> #Particion de base de datos
> data.ids<-createDataPartition(bd$ID,p=0.8,list=FALSE)# PARTIR LA BASE AL 80% DE DATOS
> bd_p<-bd[data.ids,] #cambio de variable a la base partida
> bd_v<-bd[-data.ids,] #cambio de variable a la base para la validacion (20% de datos)
> write.csv(bd_v,file = "validation_externa_data.csv") #Exporta en formato csv datos para la verificacion
> #ANALISIS EXPLORATORIO DE DATOS
> summary(bd_p) #Estatisticos generales de cada variable
      ID      Vertice      Hoja_MTN      X_ED50      Y_ED50      X_ETRS89
Min.   : 2.0 1004   : 1 Min.   :534 Min.   :429048 Min.   :4462773 Min.   :428939
1st Qu.:985.8 1005   : 1 1st Qu.:559 1st Qu.:438888 1st Qu.:4470999 1st Qu.:438779
Median :1971.0 1020   : 1 Median :559 Median :441852 Median :4475376 Median :441743
Mean   :1970.5 1021   : 1 Mean   :558 Mean   :442149 Mean   :4475415 Mean   :442040
3rd Qu.:2953.5 1022   : 1 3rd Qu.:559 3rd Qu.:445893 3rd Qu.:4479745 3rd Qu.:445783
Max.   :3939.0 1023   : 1 Max.   :582 Max.   :456064 Max.   :4490438 Max.   :455955
      (Other):3146
      Y_ETRS89      Z_Ortometr      Z_elipsoidal      ondulacion
Min.   :4462566 Min.   :543.9 Min.   :594.9 Min.   :50.97
1st Qu.:4470791 1st Qu.:620.8 1st Qu.:671.8 1st Qu.:51.01
Median :4475169 Median :657.8 Median :708.9 Median :51.02
Mean   :4475207 Mean   :653.8 Mean   :704.8 Mean   :51.03
3rd Qu.:4479538 3rd Qu.:685.3 3rd Qu.:736.3 3rd Qu.:51.04
Max.   :4490231 Max.   :762.3 Max.   :813.3 Max.   :51.09

> hist(bd_p$ondulacion, col = "blue") #devuelve histograma
> boxplot(bd_p$ondulacion, add = FALSE, col = "blue",plot = TRUE) #devuelve diagrama de caja

```

```

> mean(bd_p$ondulacion)#MEDIA
[1] 51.02516
> sd(bd_p$ondulacion)#DESVIACION ESTANDAR
[1] 0.01665814
> cv<-sd(bd_p$ondulacion)/mean(bd_p$ondulacion) # calcular el coeficiente de variacion
> cv #visualizar el coeficiente de variacion, el cual deberia ser <50%
[1] 0.0003264691
> #Para identificar si el proceso es Estacionario
> Boxplot(~ondulacion,data=bd_p,id.method="y") #devuelve diagrama de caja con otra funcion
[1] "3261" "344" "487" "3434" "3202" "441" "3237" "2798" "442" "3596" "2564" "3900" "2910" "1941" "1940"
[16] "3917" "106" "1393" "1355" "3192"
> scatterplot(ondulacion~X_ETRS89,reg.line="lm",smooth=TRUE,spread=TRUE,id.method='mahal',id.n=2,boxplots="xy",span=0.5,data=bd) #para graficar la dispersion datos para proceso estacionario en "X"
There were 30 warnings (use warnings() to see them)
> scatterplot(ondulacion~Y_ETRS89,reg.line="lm",smooth=TRUE,spread=TRUE,id.method='mahal',id.n=2,boxplots="xy",span=0.5,data=bd) #para graficar la dispersion datos para proceso estacionario en "Y"
There were 50 or more warnings (use warnings() to see the first 50)
> #Generacion de Objeto en GEOSTATS
> Ndb<-db.create(x1=bd_p$X_ETRS89,x2=bd_p$Y_ETRS89,z1=bd_p$ondulacion) #crear base de datos
There were 50 or more warnings (use warnings() to see the first 50)
> plot(Ndb,pch=21,bg="blue",col="black",title="samples location") #Vizualizacion espacial de muestras
> data.4dir.vario=vario.calc(Ndb,dir=c(0,45,90,135),lag = 500,tolang=22.5) #Generacion de variograma direccionales, tolang es el angulo de barrido/2
> data.4dir.vario #para visualizar la generacion del variograma

```

Variogram characteristics

```
=====
```

```

Number of variable(s) = 1
Number of direction(s) = 4
Space dimension = 2

```

Direction 1

```
-----
```

```

Number of lags = 10
Direction coefficients = ( 1.000 0.000)
Direction angles (degrees) = ( 0.000)
Tolerance on direction = 22.500000 (deg)
Calculation lag = 500

```

For variable 1

```
Referenced value (variance,...) = 0.000
```

| Rank | Npairs | Distance | Value |
|------|----------|----------|-------|
| 0 | 985.000 | 161.101 | 0.000 |
| 1 | 6260.000 | 537.462 | 0.000 |

| | | | |
|---|-----------|----------|-------|
| 2 | 11614.000 | 1014.927 | 0.000 |
| 3 | 16344.000 | 1512.525 | 0.000 |
| 4 | 20174.000 | 2007.694 | 0.000 |
| 5 | 23936.000 | 2507.156 | 0.000 |
| 6 | 27690.000 | 3006.019 | 0.000 |
| 7 | 31606.000 | 3504.029 | 0.000 |
| 8 | 33680.000 | 4002.740 | 0.000 |
| 9 | 36839.000 | 4503.696 | 0.000 |

Direction 2

Number of lags = 10
Direction coefficients = (0.707 0.707)
Direction angles (degrees) = (45.000)
Tolerance on direction = 22.500000 (deg)
Calculation lag = 500

For variable 1

Referenced value (variance,...) = 0.000

| Rank | Npairs | Distance | Value |
|------|-----------|----------|-------|
| 0 | 784.000 | 167.111 | 0.000 |
| 1 | 6002.000 | 539.818 | 0.000 |
| 2 | 11281.000 | 1019.822 | 0.000 |
| 3 | 16217.000 | 1512.799 | 0.000 |
| 4 | 20399.000 | 2010.319 | 0.000 |
| 5 | 24284.000 | 2508.496 | 0.000 |
| 6 | 27868.000 | 3005.567 | 0.000 |
| 7 | 31481.000 | 3504.389 | 0.000 |
| 8 | 34863.000 | 4004.793 | 0.000 |
| 9 | 38209.000 | 4503.373 | 0.000 |

Direction 3

Number of lags = 10
Direction coefficients = (0.000 1.000)
Direction angles (degrees) = (90.000)
Tolerance on direction = 22.500000 (deg)
Calculation lag = 500

For variable 1

Referenced value (variance,...) = 0.000

| Rank | Npairs | Distance | Value |
|------|-----------|----------|-------|
| 0 | 903.000 | 162.201 | 0.000 |
| 1 | 6193.000 | 538.396 | 0.000 |
| 2 | 11148.000 | 1017.345 | 0.000 |
| 3 | 15905.000 | 1512.661 | 0.000 |
| 4 | 20229.000 | 2010.266 | 0.000 |
| 5 | 24258.000 | 2505.673 | 0.000 |
| 6 | 27380.000 | 3006.054 | 0.000 |
| 7 | 31149.000 | 3504.894 | 0.000 |

| | | | |
|---|-----------|----------|-------|
| 8 | 33954.000 | 4003.489 | 0.000 |
| 9 | 36736.000 | 4504.096 | 0.000 |

Direction 4

Number of lags = 10
 Direction coefficients = (-0.707 0.707)
 Direction angles (degrees) = (135.000)
 Tolerance on direction = 22.500000 (deg)
 Calculation lag = 500

For variable 1

Referenced value (variance,...) = 0.000

| Rank | Npairs | Distance | Value |
|------|-----------|----------|-------|
| 0 | 847.000 | 164.060 | 0.000 |
| 1 | 6229.000 | 540.428 | 0.000 |
| 2 | 11213.000 | 1018.447 | 0.000 |
| 3 | 15893.000 | 1512.423 | 0.000 |
| 4 | 20193.000 | 2010.196 | 0.000 |
| 5 | 23954.000 | 2505.560 | 0.000 |
| 6 | 27133.000 | 3003.959 | 0.000 |
| 7 | 29874.000 | 3504.378 | 0.000 |
| 8 | 33110.000 | 4004.109 | 0.000 |
| 9 | 36173.000 | 4502.963 | 0.000 |

```
> plot(data.4dir.vario,title="Direccional Variograma") #para graficar la generacion del variograma
> #para plotear unicamente variogramas en un direccion especifica
> plot(data.4dir.vario,idir0=1,title="Directionalvariogram0(BLUE) - Directionalvariogram 90(GREEN)",npairdw=TRUE,npairpt=1,col="BLUE") #variograma 0° azul
> plot(data.4dir.vario,idir0=2,title="Directionalvariogram45",npairdw=TRUE,npairpt=1,add=TRUE,col="RED")#variograma 45° rojo
> plot(data.4dir.vario,idir0=3,title="Directionalvariogram90",npairdw=TRUE,npairpt=1,add=TRUE,col="GREEN") #variograma 90° verde
> plot(data.4dir.vario,idir0=4,title="Directionalvariogram135",npairdw=TRUE,npairpt=1,add=TRUE,col="BLACK")#variograma 135° negro
> #Generacion de Variograma experimental Omnidireccional
> vario_data<-vario.calc(Ndb, lag=300, tolang = 22.5, bench = 3, calcul="vg")
> vario_data #visualizacion de variograma omnidireccional
```

Variogram characteristics

=====

Number of variable(s) = 1
 Number of direction(s) = 1
 Space dimension = 2

Direction 1

Number of lags = 10
 Direction coefficients = (1.000 0.000)
 Direction angles (degrees) = (0.000)

Tolerance on direction = 22.500000 (deg)
 Slice bench = 3.000000
 Calculation lag = 300

For variable 1

Referenced value (variance,...) = 0.000

| Rank | Npairs | Distance | Value |
|------|-----------|----------|-------|
| 0 | 394.000 | 97.601 | 0.000 |
| 1 | 2433.000 | 321.411 | 0.000 |
| 2 | 4418.000 | 611.759 | 0.000 |
| 3 | 6482.000 | 905.855 | 0.000 |
| 4 | 7982.000 | 1205.417 | 0.000 |
| 5 | 9854.000 | 1504.704 | 0.000 |
| 6 | 11209.000 | 1802.319 | 0.000 |
| 7 | 12605.000 | 2101.417 | 0.000 |
| 8 | 13919.000 | 2403.462 | 0.000 |
| 9 | 15234.000 | 2702.068 | 0.000 |

```
> vario.plot(x=vario_data,npairdw=TRUE,add=FALSE,npairpt=1,title="Variograma experimental Omnidireccional") #Graficar variograma omnidireccional
```

```
> #Ajuste de Modelo para Variograma teorico
```

```
> data_model<-model.auto(vario_data,struct=melem.name(c(1,2,4,11)), auth.aniso = TRUE,col=2) #ajuste de modelo teorico
```

```
> str(data_model) #visualizar los parametros del modelo optimizado
```

```
Formal class 'model' [package "RGeostats"] with 4 slots
```

```
..@ ndim : int 2
```

```
..@ nvar : int 1
```

```
..@ basics :List of 3
```

```
...$:Formal class 'melem' [package "RGeostats"] with 7 slots
```

```
...@ vartype : chr "Exponential"
```

```
...@ sill : num [1, 1] 5.92e-05
```

```
...@ range : num 780
```

```
...@ param : num 1
```

```
...@ flag.aniso : logi FALSE
```

```
...@ aniso.rotmat: num [1:2, 1:2] 1 0 0 1
```

```
...@ aniso.coeffs: num [1:2, 1] 1 1
```

```
...$:Formal class 'melem' [package "RGeostats"] with 7 slots
```

```
...@ vartype : chr "Gaussian"
```

```
...@ sill : num [1, 1] 5.31e-05
```

```
...@ range : num 1140
```

```
...@ param : num 1
```

```
...@ flag.aniso : logi FALSE
```

```
...@ aniso.rotmat: num [1:2, 1:2] 1 0 0 1
```

```
...@ aniso.coeffs: num [1:2, 1] 1 1
```

```
...$:Formal class 'melem' [package "RGeostats"] with 7 slots
```

```
...@ vartype : chr "Stable"
```

```
...@ sill : num [1, 1] 0.000231
```

```
...@ range : num 6901
```

```
...@ param : num 2
```

```
...@ flag.aniso : logi FALSE
```

```
...@ aniso.rotmat: num [1:2, 1:2] 1 0 0 1
```

```

... ..@ aniso.coeffs: num [1:2, 1] 1 1
..@ properties: list()
> #Nube variografica
> nube_va<-cloud.calc(Ndb,dbout = NA,dirvect = NA,lagmax = 22000,varmax = 4,lagnb
= 100,varnb = 100,toldis = 0.5,tolang = NA,opt.code = 0,tolcode = 0) #generacion de nube
> db.plot(nube_va,pos.x = 1,pos.y = 2,title = "Nuve Variografica",add=FALSE,reset = TR
UE, ylim =c(0,0.04),xlim = NA) #visualiar Nube
> #Vecindad de estimacion (matriz de busqueda)
> moving.neigh<-neigh.input(ndim = 2) #por vecinos cercanos
Neighborhood type :
0 : Unique Neighborhood
1 : Bench Neighborhood
2 : Moving Neighborhood
3 : Image Neighborhood
Neighborhood type [0,3] : 2
Minimum number of points in neighborhood (Def=1) [1,NA] : 8
Maximum number of points in neighborhood [1,NA] : 16
Use sector search (Def=n) [y,n] : y
Number of sectors (Def=4) [1,NA] : 8
Maximum number of samples per sector [1,NA] : 2
Use Anisotropic search (Def=n) [y,n] : n
Maximum search distance [0.000000,NA] : 3100
> View(moving.neigh)
> #Validacion Cruzada
> data.db1<-xvalid(Ndb,data_model,moving.neigh) #Generar validacion cruzada interna
> data.db1 #visualizar

```

Data Base Characteristics

=====

Data Base Summary

File is organized as a set of isolated points

Space dimension = 2

Number of fields = 6

Maximum Number of attributes = 6

Total number of samples = 3152

Variables

Field = 1 - Name = rank - Locator = rank

Field = 2 - Name = x1 - Locator = x1

Field = 3 - Name = x2 - Locator = x2

Field = 4 - Name = z1 - Locator = NA

Field = 5 - Name = Xvalid.z1.esterr - Locator = z1

Field = 6 - Name = Xvalid.z1.stderr - Locator = z2

```

> hist(db.extract(data.db1,names=c("Xvalid.z1.esterr")),nclass = 50,main = "HISTOGRA
MA ERROR DE VALIDACION CRUIZADA",xlab = "Error validacion cruzada",col="BL
UE") #Visualizar Histogrma de errores Estimados

```

```

> db.write(data.db1,filename="errorvalidacion_interna.csv",must.noproj = TRUE,nogrid =
TRUE,flag.calcul = TRUE) #Exportar Base en formato csv (Errores de prediccion)
> #Generacion de grilla para krigin
> griddb<-db.create(flag.grid = TRUE,nx=c(540,553),x0=c(428939,4462566),dx=c(50,50)
,autoname = FALSE) #Formar grilla en funcion a columnas, filas y resolucion espacial
> print(griddb)# visualizar datos de grilla

```

Data Base Characteristics

=====

Data Base Summary

```

File is organized as a regular grid
Space dimension          = 2
Number of fields         = 3
Maximum Number of attributes = 3
Total number of samples  = 298620
Grid characteristics:
Origin : 428939.0004462566.00
Mesh   :  50.000   50.000
Number :   540     553
Angles :   0.000   0.000

```

Variables

```

Field = 1 - Name  = rank - Locator = NA
Field = 2 - Name  = x1 - Locator = x1
Field = 3 - Name  = x2 - Locator = x2
> #KRIGIN
> griddb<-kriging(Ndb,griddb,data_model,moving.neigh,radix="N.ALL") #modelo Krigin
> griddb

```

Data Base Characteristics

=====

Data Base Summary

```

File is organized as a regular grid
Space dimension          = 2
Number of fields         = 5
Maximum Number of attributes = 5
Total number of samples  = 298620
Grid characteristics:
Origin : 428939.0004462566.00
Mesh   :  50.000   50.000
Number :   540     553
Angles :   0.000   0.000

```

Variables

```

Field = 1 - Name = rank - Locator = NA
Field = 2 - Name = x1 - Locator = x1
Field = 3 - Name = x2 - Locator = x2
Field = 4 - Name = N.ALL.z1.estim - Locator = z1
Field = 5 - Name = N.ALL.z1.stdev - Locator = z2
> plot(griddb,name.image="N.ALL.z1.estim",col=topo.colors(40),title="Estimacion -All d
ata(Moving Neighborhood)") #VisiaulizaR kRIGING como raster
> plot(griddb,name.contour="N.ALL.z1.estim",nlevels=20,add=TRUE) #VisiaulizaR kRI
GING como raster con isolineas
> db.write(griddb,filename = "modelo_N_prediccion_krigin.asc",must.noproj = TRUE,nog
rid = TRUE,flag.calcul = TRUE)# Exportar formato asc grilla con kriging
> db.write(griddb,filename = "modelo_N_prediccion_krigin.csv",must.noproj = TRUE,nog
rid = TRUE,flag.calcul = TRUE)# Exportar formato csv grilla con kriging

```


Superficie de estimación de Ondulación Geoidal, generada por ArcGis para Madrid

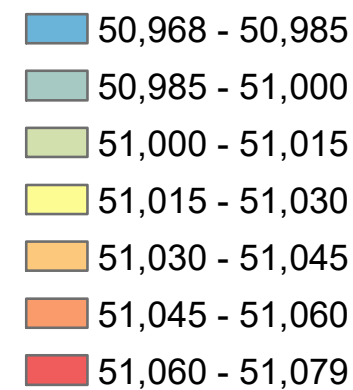
Zona de Estudio



LEYENDA

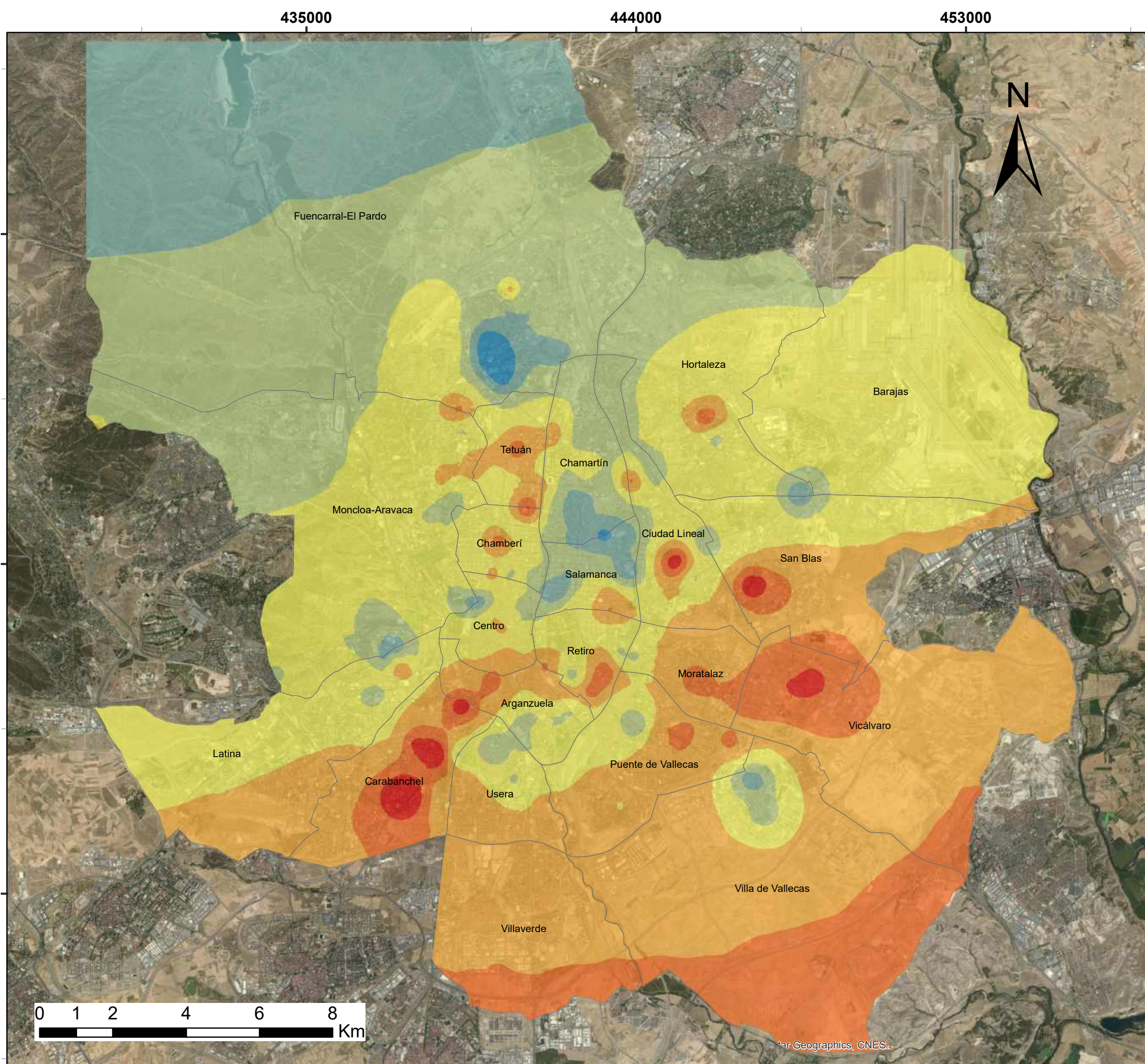
□ Límite distrital

Ondulación Geoidal (m)



UNIVERSIDAD COMPLUTENSE
MADRID

| | |
|-----------------------|----------------|
| Escala | 1:100000 |
| Institución | UCM |
| Elaborado por | Jorge Estrella |
| Sistema de Referencia | ETRS89 UTM 30N |



435000

444000

453000

Anexo N°5

Desviación de la Superficie
de estimación de Ondulación
Geoidal, generada
por ArcGis para Madrid

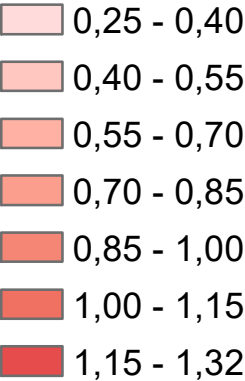
Zona de Estudio



LEYENDA

□ Límite distrital

Desv. Stn de N (cm)



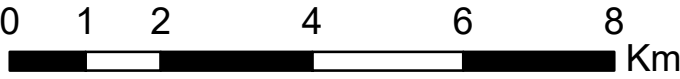
UNIVERSIDAD COMPLUTENSE
MADRID

| | |
|-----------------------|----------------|
| Escala | 1:100000 |
| Institución | UCM |
| Elaborado por | Jorge Estrella |
| Sistema de Referencia | ETRS89 UTM 30N |

4485000

4476000

4467000



Fuencarral-El Pardo

Hortaleza

Barajas

Tetuán

Chamartín

Moncloa-Aravaca

Chamberí

Ciudad Lineal

Salamanca

San Blas

Centro

Retiro

Moratalaz

Arganzuela

Vicálvaro

Latina

Carabanchel

Usera

Puente de Vallecas

Villa de Vallecas

Villaverde

Star Geographics, CNES

Superficie de estimación de Ondulación Geoidal, generada por R-Studio para Madrid

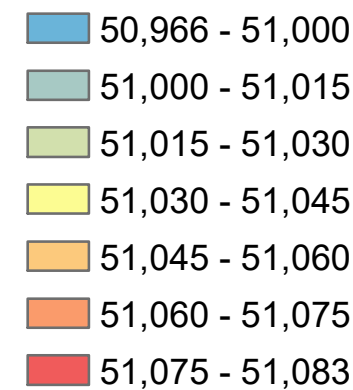
Zona de Estudio



LEYENDA

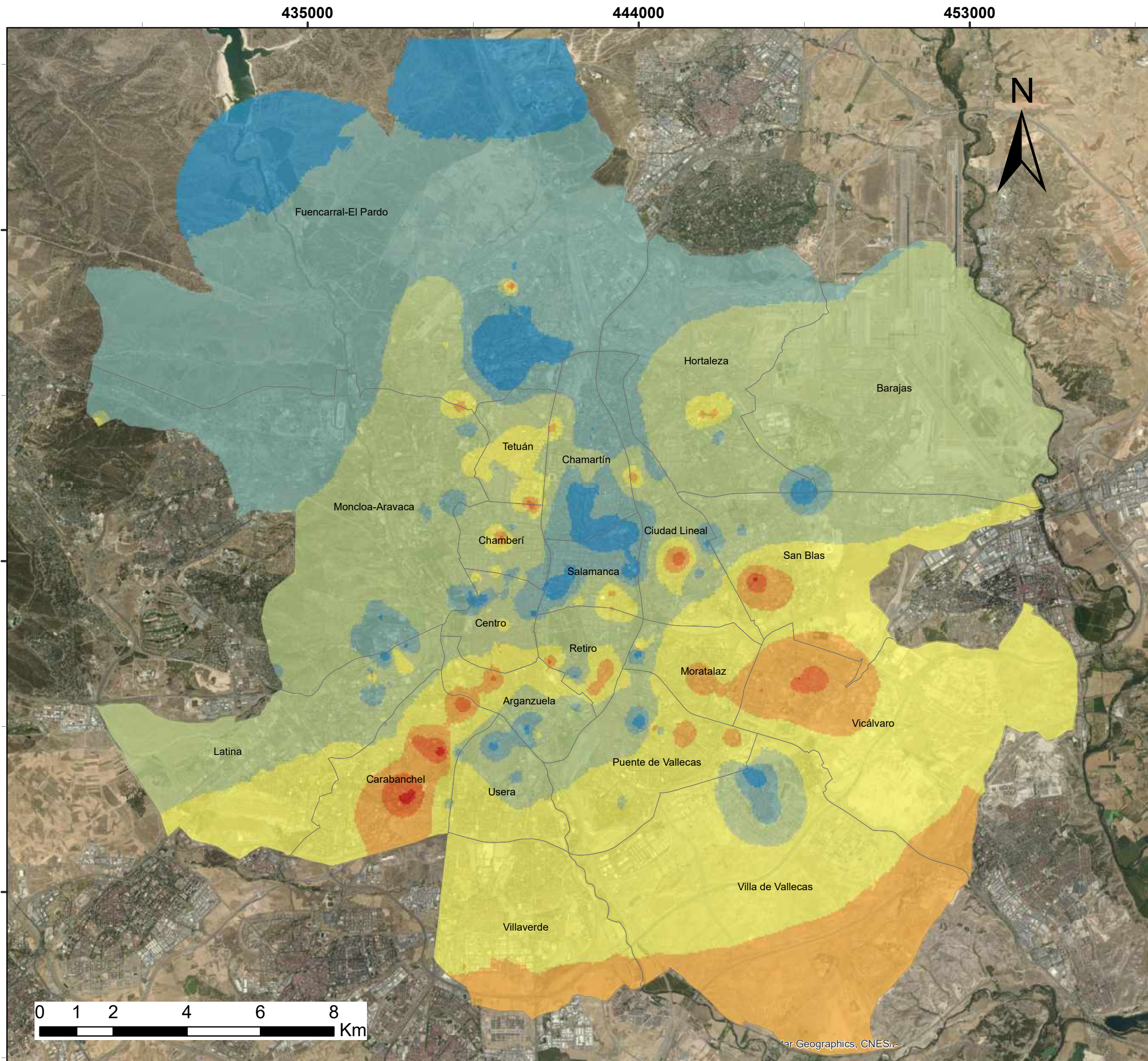
□ Límite distrital

Ondulación Geoidal (m)



UNIVERSIDAD COMPLUTENSE
MADRID

| | |
|-----------------------|----------------|
| Escala | 1:100000 |
| Institución | UCM |
| Elaborado por | Jorge Estrella |
| Sistema de Referencia | ETRS89 UTM 30N |



435000

444000

453000

Anexo N°7

Desviación de la Superficie
de estimación de Ondulación
Geoidal, generada
por R-Studio para Madrid

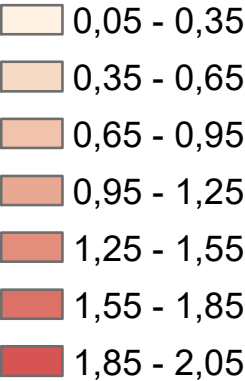
Zona de Estudio



LEYENDA

□ Límite distrital

Desv. Stn de N (cm)



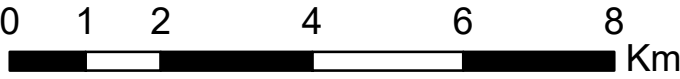
UNIVERSIDAD COMPLUTENSE
MADRID

| | |
|-----------------------|----------------|
| Escala | 1:100000 |
| Institución | UCM |
| Elaborado por | Jorge Estrella |
| Sistema de Referencia | ETRS89 UTM 30N |

4485000

4476000

4467000



Fuencarral-El Pardo

Hortaleza

Barajas

Tetuán

Chamartín

Moncloa-Aravaca

Chamberí

Ciudad Lineal

Salamanca

San Blas

Centro

Retiro

Moratalaz

Arganzuela

Vicálvaro

Latina

Carabanchel

Usera

Puente de Vallecas

Villa de Vallecas

Villaverde

Star Geographics, CNES

435000

444000

453000

Anexo N°8

Mapeo de residuales de la V.C Interna de N, generada por ArcGis para Madrid

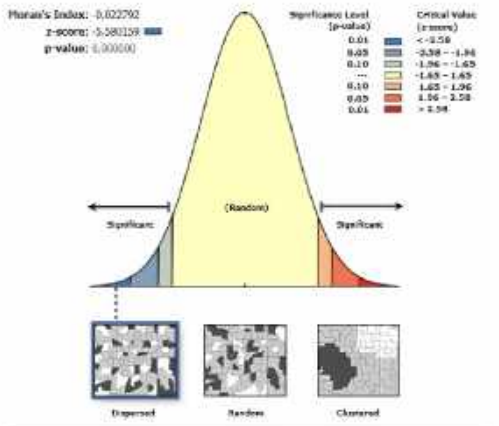
Zona de Estudio



LEYENDA

Correlación por Moran

- No significativo
- Cluster: High - High
- Outlier: High - Low
- Outlier: Low - High
- Cluster: Low - Low



Given the z-score of -5.5801588826, there is a less than 1% likelihood that this dispersed pattern could be the result of random chance.



UNIVERSIDAD COMPLUTENSE
MADRID

| | |
|-----------------------|----------------|
| Escala | 1:100000 |
| Institución | UCM |
| Elaborado por | Jorge Estrella |
| Sistema de Referencia | ETRS89 UTM 30N |

N

Fuencarral-El Pardo

Hortaleza

Barajas

Tetuán

Chamartín

Moncloa-Aravaca

Chamberí

Ciudad Lineal

San Blas

Salamanca

Centro

Retiro

Moratalaz

Vicálvaro

Latina

Carabanchel

Usera

Arganzuela

Puente de Vallecas

Villa de Vallecas

Villaverde



Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community

435000

444000

453000

Anexo N°9

Mapeo de residuales de la V.C Interna de N, generada por R-Studio para Madrid

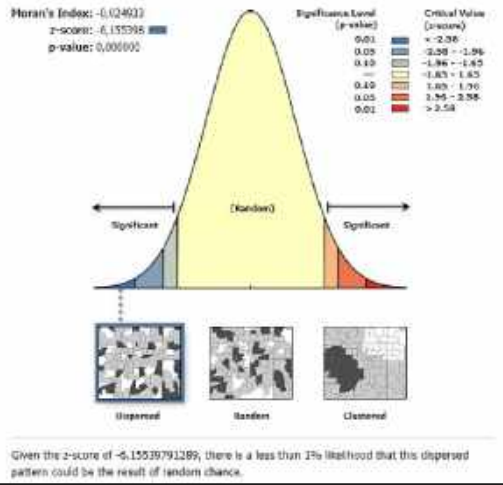
Zona de Estudio



LEYENDA

Correlación por Moran

- No significativo
- Cluster: High - High
- Outlier: High - Low
- Outlier: Low - High
- Cluster: Low - Low



UNIVERSIDAD COMPLUTENSE
MADRID

| | |
|-----------------------|----------------|
| Escala | 1:100000 |
| Institución | UCM |
| Elaborado por | Jorge Estrella |
| Sistema de Referencia | ETRS89 UTM 30N |

N



Fuencarral-El Pardo

Hortaleza

Barajas

Tetuán

Chamartín

Moncloa-Aravaca

Chamberí

Ciudad Lineal

San Blas

Salamanca

Centro

Retro

Moratalaz

Vicálvaro

Latina

Carabanchel

Usera

Puente de Vallecas

Villa de Vallecas

Villaverde



Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community

435000

444000

453000

Anexo N°10

Mapeo de residuales de la V.C Externa de N, generada por ArcGis para Madrid

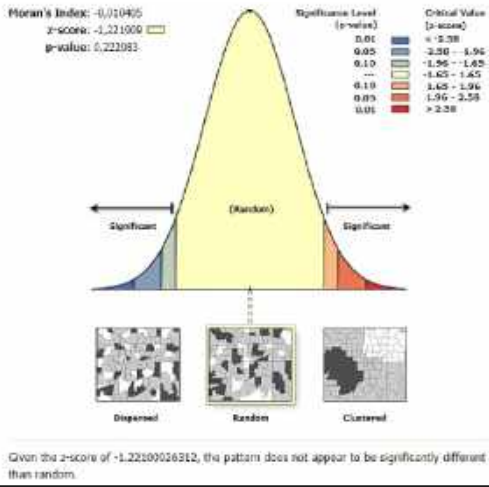
Zona de Estudio



LEYENDA

Correlación por Moran

- No significativo
- Cluster: High - High
- Outlier: High - Low
- Outlier: Low - High
- Cluster: Low - Low



Given the z-score of -1.22100026312, the pattern does not appear to be significantly different than random.



UNIVERSIDAD COMPLUTENSE
MADRID

| | |
|-----------------------|----------------|
| Escala | 1:100000 |
| Institución | UCM |
| Elaborado por | Jorge Estrella |
| Sistema de Referencia | ETRS89 UTM 30N |

N



Fuencarral-El Pardo

Hortaleza

Barajas

Tetuán

Chamartín

Moncloa-Aravaca

Chamberí

Ciudad Lineal

San Blas

Salamanca

Centro

Retiro

Moratalaz

Arganzuela

Vicálvaro

Latina

Carabanchel

Usera

Puente de Vallecas

Villa de Vallecas

Villaverde



Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community

435000

444000

453000

Anexo N°11

Mapeo de residuales de la
V.C Externa de N, generada
por R-Studio para Madrid

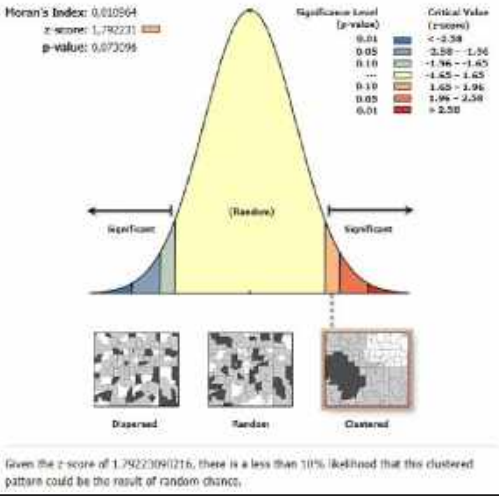
Zona de Estudio



LEYENDA

Correlación por Moran

- No significativo
- Cluster: High - High
- Outlier: High - Low
- Outlier: Low - High
- Cluster: Low - Low

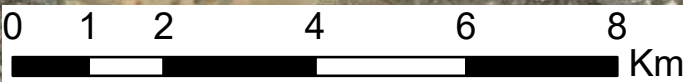


Given the z score of 1.79223091216, there is a less than 10% likelihood that this clustered pattern could be the result of random chance.



UNIVERSIDAD COMPLUTENSE
MADRID

| | |
|-----------------------|----------------|
| Escala | 1:100000 |
| Institución | UCM |
| Elaborado por | Jorge Estrella |
| Sistema de Referencia | ETRS89 UTM 30N |



Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community